VISIT…

LANZAROTE Caliente.COM



李葆嘉 著

# 理论语言学

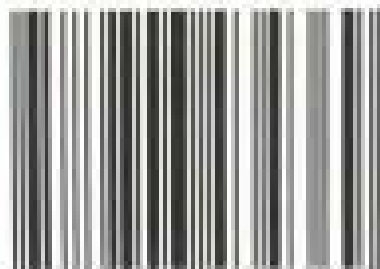
人文与科学的双重精神

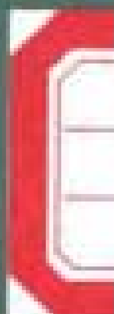
江苏古籍出版社



理论语言学

人文与科学的双重精神


ISBN 7-80643-436-4
9 787806 434369 >



ISBN 7-80643-436-4
H·17 定价：18.80元

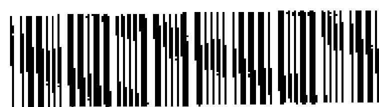
李葆嘉 著



# 理论语言学

## 人文与科学的双重精神

江苏古籍出版社



图书在版编目(CIP)数据

理论语言学:人文与科学的双重精神/李葆嘉著.
南京:江苏古籍出版社,2001.8
(随园新积丛书)
ISBN 7-80643-436-4

Ⅰ.理… Ⅱ.李… Ⅲ.语言学 Ⅳ.H0

中国版本图书馆 CIP 数据核字(2001)第 053456 号

**理论语言学:人文与科学的双重精神**

著 作 者 李葆嘉
责任编辑 冯保善

出版发行 江苏古籍出版社
发行部电话 025—3223462
社 址 南京市中央路 165 号 邮编 210009
经 销 江苏省新华书店

印 刷 者 淮阴新华印刷厂
开 本 大 32
印 张 11.75
印 数 1—2000 册
字 数 272 千字
版 次 2001 年 9 月第 1 版第 1 次印刷
标准书号 ISBN 7-80643-436-4/H·17
定 价 18.80 元

(江苏古籍版图书凡印装错误可向承印厂调换)

# 总　序

随园，是清乾隆时期著名文学家袁枚的故居。南京师范大学，恰好坐落在随园旧址大略所在的那一片山丘及其附近的旷地。校区内古木参天，芳草茵地，鸣禽时变，檐翼欲飞，风景之清幽美丽，俨然是一座道地的南国园林。因此，许多年来人们习惯上也将南师校园称作"随园"。

南京师范大学文学院中国古代文学学科，是1981年国务院学位委员会批准的全国首批博士学位点和硕士学位点之一，又是我院"博士后流动站"和"国家文科人才培养和科学研究基地"的骨干学科，国家"211工程"重点建设项目"中国传统文学与现代文化"的带头学科。本学科教师的教学和研究方向，涵括古代文学、文献学、古代文论等各个门类，从先秦直至晚清等各个时代，诗文、词曲、戏剧、小说等各种文体，覆盖面甚广。其中唐代文史研究、词学研究、明清小说研究等在国内外学术界具有较大的影响。

建国以来，本学科曾经拥有唐圭璋、孙望、段熙仲、诸祖耿、吴调公等著名学者，出版过《全宋词》、《全金元词》、《词话丛编》、《元次山年谱》、《水经注疏校点》、《战国策集注汇考》、《李商隐研究》等高质量的著述。他们不仅以自己严谨的学风、精深的知识为本学科奠定了坚实的基础，而且亲手培养了几代教学与科研人才，催生了一批又一批新的教学与科研成果，诚所谓"新笋已

成堂前竹，落花都入燕巢泥"。

为了不断地向学术界介绍我院中国古代文学学科的最新研究成果，增加整个中国古代文学研究的学术积累，我们决定精选本学科教师在古代文学之文献与理论研究等方面的优秀著作，编为这套《随园薪积丛书》。所谓"薪积"也者，包括两个层面的含义：

第一，有取于《史记·汲郑列传》"积薪"而"后来者居上"之喻，希望我们的研究成果能够不断提高学术品位，提高学术价值。

第二，有取于《庄子·养生主》"指穷于为薪，火传也，不知其尽也"之喻。教师是"薪"，学术是"火"，"薪"尽而"火"传。我们的物质生命不免于毁灭，但精神的生命却可望凭藉学术著作的流传而达到永恒。

有冀于此，我们期待本学科的同仁将自己的"随园魂"融入这套丛书，使她真正成为我们精神生命的累积！

《随园薪积丛书》编委会

# 自序　语言是人类的精神家园

根据人类学家的见解——现代语言造就了现代人类和现代文化，这一“现代”，不是指工业革命以后的“现代”或20世纪的“现代”，而是指人类进化史上的“现代”，即从距今3.5万年之际以人类语言成熟为标志而开始的“现代”。从这个意义上来说，所谓人性就是人的“语言性”，所谓世界就是“语言的世界”，所谓世界观首先表现为“语言世界观”。

**语言符号是人类特有的认知符号系统**

语言既是人类与自然的纽带，又是个人与人群的桥梁。语言的成熟是历史上的第一次信息革命，提供了人类组成社会的必要条件。文字的发明是历史上的第二次信息革命，成为人类进入文明阶段的明确标志。人们通过语言符号来建构对象世界，对象世界在语言系统中才具有精神价值，人所认识的世界就是语言中的世界。人类的精神世界可以分为心智机制、语言符号和文化形态三个层面。在源于动物性感受及植根于经验基础之中形成的人类心智机制，成为初民语言符号化认知的出发点，而语言符号产生以后又作用于心智机制，并形成心智和语言的互塑。正是依据语言符号系统，人类才从不可知的混沌得以进入可认知的清晰。同样，表现为语义结构的感知就是显意识，反之则是潜意识。人类借助语言文字符号系统，在原始巫术和神化传说的史前意识基础上建构起宗教、艺术、道德、制度和科学

的意识形态大厦。物化的语言文字,作为主观与客观的枢纽和心智与文化的中介,从而呈现出人类历史长河中的纵横型多边互动。

**语言文字是人类一切学科的底座**

文化是人类对生存环境的适应。一方面,语言文字是认知符号或本原文化的本体,另一方面,语言文字又是知识底座或派生文化的载体。从学科发生来说,创造文字的本初语言学是人类最古老的学科,比西方所谓"一切学科之母"的古希腊哲学,实为借用语言学音素论模式处理自然神话的"希腊奇迹"要早得多。从学科功能来说,"文字者,经艺之本,王政之始",中国历代统治者都把语文建设作为文治之盛举,经学居"四部"之首而小学是经学之根。从学科关系来说,语言学是人文社会学科与自然科学的桥梁,各学科的理论和方法相互渗透。虽然语言学的谱系树和生物学进化论的创立先后难定,但结构主义和生成语法,对20世纪诸多学科的巨大影响则众所周知。至于语言论美学的狂飙突进百年,语言论哲学成为现当代西方哲学的主流,更是学术史上不争的事实。

**语言结构模式是其他文化模式的基础**

在洪堡特的"语言世界观说"和萨丕尔—沃尔夫假说基础之上,我提出"语言文化结构模式同构论"。主要观点是:1.语言结构模式及其自发投射的文字结构模式即本原认知结构模式。语言结构特征决定了语音分析模式,屈折语语音的继起性产生了音素论,词根语语音的共起性导致了音节论。欧洲自然哲学的原子论模式与音素论同构,中国人伦哲学的玄牝象模式与音节论同构。2.语言结构模式与信仰、哲学、文化、经济、政治结构模式等之间同样具有相关性,某一类型中的子文化模式只是这些模式在衍生文化领域内的历史性延伸。所谓相关性,一是指结

构模式之间的制约性，不是任何细节都面面俱到；二是指结构模式之间的互动性或互塑性，不是仅仅语言文字模式的单控，而是各结构模式、意识形态之间的互动影响。3. 人类对世界的认知，除了语言符号这一主要方式，还有其他感知方式与之互相作用、互相渗透。影响人类文明发展进程的因素是多样的，但是在旧石器时期晚期与新石器—金属时期最重要的客观因素是生存环境。在人类文明的早期，神经机制和生存环境的交汇作用决定了某一群体的认知经验框架与集体无意识，决定了该语言结构模式和其他相关文化模式的结构趋同性。既然语言文字学是一切学科之母。唯有以语言文字模式为线索，才可能贯通原始崇拜、神话学、宗教学、历史学、哲学、思维学、逻辑学、社会学、政治学、心理学、自然科学和艺术学等，以揭开东西方语言文化哲学模式的生命密码。

**当代语言学已经成为一门领先科学**

20 世纪 50 年代以后，在理论方法交叉而形成繁荣的边缘语言学家族的同时，语言学与计算机科学的结合则产生了计算语言学或语言工程学。随着二者的天然结合，语言学在当代学科体系中成为领先科学。语言学的至关紧要性，可以第五代计算机因未能穿越语义形式化的瓶颈而不得不搁浅来证明。21 世纪的“认知神经科学领域”和“信息技术领域”，分别提出“探索意识、思维活动的本质”和“计算机向智能化方向发展”的重大任务，然而没有语言学的根本突破则不可能实现。

**语义语法学是人类普遍语法学理论**

自然语言的计算机处理成为当代国际语言学研究的共同目标。认知语义性是人类语言的共同本质属性。只要人们探索语言本质的好奇永远存在，只要“人脑语言移植于电脑”的目标不变，就必须穿越语义结构形式化研究的沼泽地，因此有必要在前

修时贤的基础上,建构与绵延了2500年的语形语法学相区别、以人类语言共性为基础的语义语法学(semantic grammar)理论。语义语法学包括"语义语法学的理论建构"和"语义语法学的应用研究"两个方面,具有理论研究和应用研究的双重目的。其理论价值在于建构符合人类语言共性和语法真实本体的语法学理论,其应用价值在于启动汉语语义结构体系的形式化研究,为语言信息处理提供可编程文本。

**语义研究是语言基因图谱分析工程**

语义结构体系的形式化研究,可以称之为"语言基因图谱分析工程",一个词等于一个细胞,语言基因等于元语言系统。通过30,000个词的解释性义证分析和关联性要素分析,可以完成语言基因图谱分析工程。语言学家务必彻底消除"语义泥潭"的畏缩情结,与人类基因图谱中的30亿个碱基的识别鉴定相比,与人类基因数量类似的某种语言中的3万个词语的分析描写任务并不过于繁重,因此语义语法学的研究课题才形象地喻为"语言基因图谱解析工程"。语义结构的研究,是语言学家而且也仅是语言学家才能承担的根本任务。语义语法学的创建,既为国际语言学和汉语语法学提供了崭新的理论,又为语言信息科技工程找到了可行的思路。

与拘泥于"窄化语法"而难以跳出"语形框架"不同,语义语法学抓住语言的真实本体,试图以语义研究为关键贯通心理语言学、功能语言学、认知语言学、数理语言学和计算语言学等。以往的语法学研究,除了执著于语义结构表层投影的语形研究以外,还有一个根本缺憾,就是不具备自然科学研究那样的实验手段,而难免经验性、臆断性、游移性和非系统性。语言基因图谱分析工程以计算机操作为验证手段,使语言法则的研究具有自然科学的可证伪性,在语义本位的基础上促成21世纪的语言

革命。以语义语法学理论为指导的语言基因图谱分析工程，具有科学性、创新性、操作性、验证性和扩展性，研究成果无疑具有预期性和技术化前景。

**语言科技是21世纪的新思维**

依据建立与网络文化、智力经济密切相关的新学科的当代要求，我进一步提出"语言科学与语言技术"新概念或新思维。"语言科技"的内涵是以理论语言学和应用语言学研究为枢纽，以计算机科学和现代教育技术为延伸，贯通文理工以实现研究成果的技术化。"语言科技"的外延是语言信息科技、语言教学科技和语言研究科技。语言信息科技是语言学和计算机科学的结合。培养语言信息科技人才的目标，是实现汉语的计算机分析和生成。语言教学科技是语言学与现代教育技术的结合。作为大国形象标志的对外汉语教学，与一个国家未来休戚相关。无论是对外汉语教学，还是母语汉语教学，都必须将现代教学内容和现代教学手段相结合。培养语言教学科技人才的目标，是实现语言教学的媒体多维化和网络远程化。语言研究科技是语言学与现代化研究手段的结合。这一结合包括电子语料库检索系统研制、电子词典和词库系统研制，语言研究的分析性、统计性、比较性和实验性软件的开发。培养语言研究科技人才的目标，是实现语言研究自身的现代化。

南京师范大学正在组建具有"前卫性"的语言科技系，并着手建立一个包括科研功能、教学功能、开发功能和交流功能的语言科技实验中心，其目标在于形成"本科生·研究生·博士后"的培养体系，建构"教学型·科研型·技术型"的综合模式。语言科技系人才培养的总体目标定位在：面向社会的迫切需求，立足语言科技领域，放眼相关科技前沿，以语言学与计算机科学、语言教学与现代教育技术、语言研究与现代化手段相衔接为导向，培

养多层次复合型创新人才,为人类科技进步和社会发展做出贡献。

在新世纪元年,面对世界语言学的古往今来感慨万千,回顾和前瞻之余可以断言——19世纪是历史语言学的世纪,20世纪是语形语言学的世纪,而21世纪定是语义语言学的世纪。

本书的专题"语言的本质共性和语义语法学理论",是江苏省哲学社会科学"十五"规划基金课题"面向语言工程的语义语法学研究"的阶段性研究成果之一。本书的出版,得到南京师范大学文学院"211工程"重点学科建设项目出版基金的资助。

是为序。

2001年8月于苏州路寓所

# 目　　录


自序　语言是人类的精神家园 ………………………………（1）

第一章　人类语言的起源和演变模式 …………………………（1）

第一节　现代人类语言起源的科学思考 ……………………（1）

语言起源探索中的“语言”界定——现代语言起源的化石解剖学思考——现代语言起源的分子遗传学思考——人类思维交际符号演变的连续统——现代语言造就了现代人类和现代文化

第二节　超越谱系树模式的语言关系类型学 ………（19）

语言关系类型历史演变的差异性——世界语言关系类型学的四分法——语言结构类型历史演变的连续统——人类语言文化的世界化进程

第三节　中国混合语的研究现状与理论探索 ………（40）

混合语研究的历史回溯——民族混合语研究述评——历史混合语研究述评——混合语研究的价值与方法

第二章　人类文字的起源和演变规律 ………………………（58）

第一节　人类文字起源的多元发生论 ……………………（58）

文字起源的神创说与学术性假说——原始记事方法与文字原理的成熟——从句符文字的特点透视文字发生——苏美尔文字起源于粘土标志说——文字发生

的社会背景和触发需求——文明的多样性与文字的发生多元论

第二节 自源文字的发展和借源文字的突变 ……… (76)

自源文字并未发展成为拼音文字——自源文字在历史上的命运——在借用突变中产生拼音文字——以汉字为本源的借源文字——东西方文字的不同是类型的不同

第三节 语言类型与文字类型的制约关系 ………… (88)

自源文字类型的自在选择——借源文字类型的自觉选择——语言类型对文字类型的刚性和弹性制约

**第三章 语言符号的任意性和可论证性** ………………… (103)

第一节 荀子的王者制名论与约定俗成说………… (103)

荀子约定俗成说的通行曲解——荀子正名思想的王者制名论——荀子正名思想的约定俗成说——荀子约定俗成说曲解的溯源

第二节 索绪尔符号任意性原则的失误与复归…… (112)

对符号任意性原则的质疑——任意性原则证明过程中的三个失误——绝对相对任意性划分中的三步复归——任意性原则的展开反证了可论证性

第三节 语言符号的论证模式及其价值…………… (126)

任意性的含义及其对立性概念——单纯性符号的两种论证模式——根词推断论证性的类比证明——语言符号论证性原则的价值

**第四章 静态语言学的来源和心理索绪尔** ……………… (135)

第一节 静态语言学的三个直接来源……………… (135)

《教程》索绪尔和学术索绪尔——德克海姆的社会理论:索绪尔语言论的哲学基础——博杜恩的语言理

论；索绪尔语言论的理论框架——惠特尼的符号学说；索绪尔语言论的符号原则——当代世界语言学对结构主义的全面超越
第二节　语言学大师之谜和心理索绪尔…………（155）
大师索绪尔身后留下的疑团——中老年索绪尔的神经症缠绕——青年索绪尔的首创权情结——探索者索绪尔的人生悲怆——打开索绪尔的"黑匣子"
**第五章　汉语史理论模式与东亚南洋语言文化圈**………（166）
第一节　论汉语史研究的理论模式…………………（166）
汉语史研究理论模式述评——原始华夏语三元混成发生论——秦汉以后汉语推移发展论
第二节　中国语的历史和历史的中国语……………（187）
三大考古文化系统与三大太古语言区域——华夏语的混成发生与类型的动态转变——南耕北牧生态环境中的汉语嬗变轨迹——明清官话的市民社会内涵及其基础方言——七千年中国语历史演变的来龙去脉
第三节　东亚南洋语言历史文化圈假说……………（213）
东亚南洋语言关系假说的歧异——东亚南洋语言文化圈的史前渊源——东亚南洋语言文化圈的变迁鸟瞰——华夏汉语和周边语言同源异流现象阐释
**第六章　语言的本质共性与语义语法学理论**…………（248）
第一节　语法形态范畴的本质是语义认知范畴……（248）
人类语言符号与范畴化认知能力——事物的原始认知范畴或语义类别——事件的原始认知范畴或语义类别——空间和状态的原始认知范畴或语义类别——认知语义范畴化的基本特点——语法形态范畴的消亡和继起的语法化——语言的音义二分系统以及传统语

法学的位置
第二节 面向信息处理的语义语法学理论…………（279）
语言学流派和语法学目标的嬗变——人类语言的本质共性是语义性——语义语法学的语法观——语义语法学的语域观——语义语法学的研究体系
**第七章 东西方文化的本原差别及其分野**………………（315）
现代地理学分界与文化中心主义的视角——地中海文化圈与东亚南洋文化圈——西方神质宗教文化和中国人伦宗法文化——西方屈折语言文化和中国词根语言文化——西方音素论文化和中国音节论文化——西方思辨逻辑文化与中国顿悟意象文化——希腊原子论模式与音素论模式的同构性——中国玄牝象模式和音节论模式的同构性——西方演绎科学文化与中国经验技术文化——西方复式焦点艺术文化和中国单式散点艺术文化——文化类型划分的标准与东西方文化的分野——音素论和音节论在人类文化史上的地位
**跋 走出人文主义与科学主义对立的怪圈**………………（357）

# 第一章　人类语言的起源和演变模式

## 第一节　现代人类语言起源的科学思考

### 一、语言起源探索中的"语言"界定

关于语言的起源，现当代的欧美有两种对立的观点：一种是突发型模式，认为语言是现代人的独有特征，语言能力是随着脑子增大而在晚近迅速出现的一种能力；一种是连续型模式，认为语言随着人属的进化而开始，通过作用于各种认识能力的自然选择而逐步成熟。所谓现代人类或新人，是指我们这些具有意识表现力、艺术想象力和语言交际力的人类，所谓现代语言就是指现代人类所使用的语言。

在语言出现以前，尽管远古人类使用的交际手段可以看作广义语言的一部分，但是一旦将这种在动物性感受与交换信息手段基础之上发展起来的非有声分节语言与有声分节语言混为一谈，语言起源的讨论则含混不清。为了避免把语言的起源和人类符号的起源相混淆，首先有必要对"语言"做出明确的界定。语言具有三个显著特点：1.语言的人类性。虽然人类符号交际能力产生于和其他动物有着相似点的原始交换手段，但语言是现代人特有的信息交换符号，就其抽象性、孳生性和开放性而

音，与动物性信息交换方式有着本质区别。所谓“动物的语言”仅是“语言”一词的泛化。2.语言的有声性。语言是语音构成的符号系统，用“语言”指称人类的“非语音符号交际工具”是引申用法。3.语言的分节性。表达一定语义并可分为不同音节的语音，不但与动物嚎叫的连续性发音有着本质区别，而且也与早期人类的非音节性语音有着质的区别。只有首先讨论清楚“有声分节语言的起源”，然后才有可能在“非有声分节语言”与“有声分节语言”之间加以连续性沟通，以推测人类思维交际符号的演变连续统。

## 二、现代语言起源的化石解剖学思考

虽然语言起源于劳动说揭示了语言与社会劳动的依存条件，语言与思维同时起源说揭出了语言与神经思维的交互作用，但是这些思辨性论断除了诱导人们把语言的诞生时间任意前推以外，还有一个共同的盲点：忽视了语言出现的关键条件是发音器官的成熟。化石智人的发音器官是否能够发出清晰的元音，才是语言诞生的惟一准绳。当20世纪70年代的美国学者将语言起源研究的目光投向探索化石智人发音器官的功能状态时，就预示着这一研究的由思辨性转向实证性。

现代人之所以能够发出清晰的音节和不同的音素，这是由于喉部位置较低，在声带之上存在一个较大的音室——咽腔。除了现代人，所有哺乳动物喉部都位于喉咙高处，在能呼吸和饮水并举的同时限制了发音功能。儿童的体质发育史折射了人类的体质进化史。为适应吸奶和呼吸并举的需要，婴儿和哺乳动物的喉部位置一样，但在18个月以后开始下移，14岁时到达成年人的位置。如果能够确定喉部在人类进化不同阶段的高低位置，就可以对这些个体的发音状况或语言能力做出若干推定。

虽然发音器官软组织不可能作为化石保存下来，但是依据头骨底部哺乳动物平直形而现代人拱形这一差别，解剖学家认为可以利用化石头骨底部的形状来模拟发音器官的构造，即依据颅底曲度来推断人类祖先的发音功能[①]。

美国科学家选择了生存在距今4.5万年前的一个成年男子的化石。解剖学家克里林(E. Crelin)制作了这位古典尼人的模拟声道，语言学家利伯曼(P. Lieberman)利用计算机测定了通过这一声道模型可能发出的语音。结果表明，尼人不可能发出清晰的、分节的语音，其中包括像[a][i][u]这样的元音，其原因在于口腔与喉部直接相连，没有现代人的咽腔或咽腔过于狭窄，由此推论尼人没有现代人的这种有声分节语言[②]。利伯曼认为"尼人文化的普遍水平表明，他们的有限发音能力看来已充分加以利用，因此他们具有某种类型的语言"，而苏俄学者帕诺乌(E. H. Панов)却认为"从尼人的物质和精神文化是那样复杂多样这一点来加以判断，他们的语言应是相当发达的了"[③]。利伯曼的"某种类型的语言"，也许是含混语音伴随性的手势语；帕诺乌的"物质和精神文化的复杂多样性必然有相当发达的语言"，其理由并不充分。语音天然为语言，但语言并非天然是语音，远古复杂多样的文化可以借助"非语音符号"来运作。语言的产生并不是需要或想要就可以"脱口而出"，必须有赖于发音器官的成熟。

---

① [美]理查德·利基《人类的起源》(吴汝康等译)，上海科技出版社1995年。

② 唐建《语言的起源和思维的起源》，《语文论丛》(1)，上海教育出版社1981年。[美]P. Lieberman: *On the Origins of Language: An Introduction to the Evolution of Human Speech*, New York, 1975.

③ [俄]帕诺乌《尼安德特人会说话吗?》(于润译)，《国外语言学》1983年第1期。

即使如此，尼人的物质和精神文化的复杂多样性与声道的相当原始性，仍然使人们陷入某种困惑和茫然。不得不重新思考：颅底曲度大是否就意味着具有充分发达的语言？莱特曼(J. Laitman)认为，发现充分弯曲的颅底的最早时间是在距今30—40万年的远古智人中，可是依据当时的"文化"状况，这些智人似乎并不可能具有发达的语言。尼人的颅底曲度甚至不如早几十万年前的别的远古智人，这种进化上的退化似乎是不可能的①。虽然人类学家认为更为可能的答案是：尼人的头骨解剖学结构特征是对寒冷气候的一种适应，这种适应可能影响了颅底形状而又不致于明显降低发音能力，但是这一解释仍然明显缺少说服力。

既然尼人的颅底曲度较平直并非意味着他们的退化，那么是否可以对"颅底曲度与咽腔成熟度的相关性"这一预设提出质疑。根据英国与荷兰科学家的报告，通过检查提供平衡感的内耳机制，发现在10万年前进化为现代人的祖先才是第一种完全直立步行的灵长类动物②。这似乎在提醒：发音器官的成熟——咽腔与口腔的分离——与人类祖先的完全直立姿态有关。四肢俯行的动物，口腔势必直接插人喉部；完全直立的人类，才在口腔与喉部之间形成了一个弯道。与其说咽腔的形成与颅底曲度具有相关性，不如说与直立程度具有相关性。正是由于直立，才有了喉部与口腔之间的弯道；正是由于有了弯道咽腔，才有了对气流缓冲的控制能力，语音的分节才成为可能。现代人约有50种音素而猿类大约只有12种，在具有一定数量的音素

① ［美］理查德·利基《人类的起源》，101页。

② 带良编译《人类祖先何时开始直立行走，耳朵泄露了其中秘密》，《世界科技译报》1994年7月30日

以后，决定现代语言形成的关键并不是音素的多少而是语音的能否分节。只有语音的分节，才为音素的若干组合提供了可能。如果颅底曲度与咽腔形成没有必然相关性而咽腔的成熟度与身体直立度具有相关性，那么仅仅依据颅底曲度模拟声道以测定其发音功能推出的结论就不可靠。也许随着新人化石的不断发现，不是对尼人化石而是对新人化石进行声道模拟的连续性序列研究，对现代人发音器官的成熟才可能有明晰的看法。

语言的形成除了发音器官的成熟，还需要相应的神经系统的成熟。美国神经生理学家史密斯(C. Smith)认为，一直到10万年前，人类祖先的神经生理状况和别的动物没有太大差别，那些较为复杂的狩猎活动、地区迁移、社会组织、工具制造和环境控制等，都是由较为简单的符号性学习、记录、储存和革新的神经过程执行的。然后，一系列的前语言在某一方面的预先适应，与大脑进化到关键体积的相一致导致了语言的突然发明[①]。虽然依据颅底曲度模拟尼人声道以测定其发音功能的结果大可怀疑，但人类学家既然普遍同意在3.5万年前已经灭绝的尼人不是现代人类的祖先，那么根据现代人类的完全直立在距今10万年前和神经生理状况的突变在距今10万年时，就可以推断人类有声分节语言的出现不会早于这之前。怀特(R. White)提出，早于10万年前的人类各种活动的证据显示了“完全缺乏会被现代人认为语言的东西”。虽然解剖学上的现代人这时已经出现，但是还没有发明文化涵义的语言，还不是心智上的现代人。直

① [美]Curtis Smith: *Ancestral Voices: Language and the Evolution of Human Consciousness*, New Jersey: Prentice-Hall, 1985；杜诗春《语言起源问题新探——〈祖先的声音〉评介》,《国外语言学》1993年第1期。

到3.5万年前，我们的祖先才掌握了现代语言和文化[①]。

## 三、现代语言起源的分子遗传学思考

达尔文在《物种起源》中预见："如果我们拥有一份完备的人类谱系图表，那么人类种群形成的序列将为全世界所使用的各种不同的语言提供最好的分类"[②]。现代语言的起源和流播，与现代人的起源和迁徙密切相关。依据区域进化论即各地新人单独起源的多源论，必然推定现代人种的多样论和现代语言的多源论。与之相反，持有换代论或伊甸园论的人类学家认为，现代人源于东非大草原的新人，然后流向全世界。新人起源的单源论也就蕴涵了现代人种的同一论和语言起源的单源论。

区域进化论的证据主要来自古人类化石解剖学的特征，通过古人类化石和当地现代人的解剖学特征的比较而证明二者具有相承关系。必须认识到：1. 某地出土的古人化石与当地现代人之间没有必然联系。出土古人化石这一事实仅仅表明这一古人曾在这一地区生活，并不能证明就是当地人的祖先。除非具有可靠的连续性材料，否则这一推测就是"结论溢出发现"。2. 仅仅依据解剖性状的相似性并不能证明区域进化。某些性状既不限于某一区域，要证明某些性状具有区域连续性的参考点又太少。虽然铲形门牙在中国人中间占80～90%，在非洲占10%以上，在欧洲占10%以下，但是这毕竟不是只有东亚人种才具有的惟一特征。伦敦自然博物馆的斯特林格尼(H. Stringer)指出，在颜面形状、颧骨构造和铲形门牙等方面的这些性状并不局

① [美]理查德·利基《人类的起源》，97页。

② 转译自[美]王士元：*The Ancestry of Chinese: Retrospect and Prospect*，原载美国《中国语言学报》专题论丛系列之八 *The Ancestry of Chinese Language*，1995年。

限于亚洲北部,所以不能用作区域连续性的证据。3.人类解剖学性状研究的相似比较法,得出的结论只具有或然性,而不具有必然性。

与之相反,遗传基因的研究为换代论提供了佐证:1.分子生物学发现细胞质中的线粒体也含有遗传物质DNA;2.现代生殖学也已证明在高等动物的受精过程中,精子中的线粒体DNA是不能够进入受精卵的,人类细胞的线粒体DNA都是来自母亲;3.既然线粒体DNA属于严格的母系遗传,那么如果能够证实现代人类的线粒体DNA是相同的,就可以推断他们来源于同一个母系。1987年,美国伯克莱加州大学威尔逊(A. Wilson)小组的科学家吕贝卡·卡门(Rebecca L. Cam)等人,选择了来自非洲、欧洲、中东、亚洲以及新几内亚和澳大利亚土著妇女147人生产婴儿的胎盘,进行不同种族胎盘的线粒体DNA研究。发现人类线粒体DNA基本相同,平均差异率只有0.32%左右。由此推定现代各民族居民的线粒体DNA,最终都是一个共同女性祖先的遗传。迄今为止,接受测试的人数已经达到4000多人,生物学家宣称没有发现具有古老来源的不同于现代人的线粒体类型。由此美国生物学家提出现代人源于15～12万年生活在东非的同一个聚落的共同祖先,在距今大约10万年之后向全球扩散开来,与各地原有的直立人并不相干的"夏娃理论"。支撑这一理论的还有"新人化石·石器文化"的考古学证据。在非洲发现了距今10万年前的新人化石和距今大约9～7万年新人所建立的晚期旧石器文化。在非欧亚通道的中东地区,发现了距今大约9万年左右的卡夫扎洞穴新人化石[①]。根据夏娃理论,在夏娃后代迁往各地的过程中,虽然古典尼人和山

① [美]理查德·利基《人类的起源》,69页。

顶洞人这些晚期智人已在欧亚各地生息繁衍，但是夏娃后代并没有与当地原人混合交融而是完全取代了他们。否则那些欧亚智人就会将与夏娃人种不同的线粒体 DNA 遗传下来，现代人中就会具有不同种类的线粒体 DNA。然而研究思路决定了对实验结果的解释，为了判断正确与否，有必要向夏娃理论挑战。

可以采取完全相反的思路：预设欧亚原人与夏娃新人的线粒体 DNA 本来就相同，而现代人类是她们的共同后代。根据高等动物线粒体 DNA 母系遗传论，如果夏娃时代分布在欧亚非的所有智人都起源于同一古老的猿人——能人——古人——早期智人的母系，如果这些祖先都属于现代生殖学研究所界定的“高等动物”，那么通过更早迁徙而分布在欧亚非并由此进化而来的 15—10 万年之间的人类的线粒体 DNA 不也就应当相同吗？换而言之，现代人类线粒体 DNA 的相同性可能源于更为远古的人类线粒体 DNA 的同源性。夏娃理论的研究者认为，当时也许有几千男女同“夏娃”生活在一起，但是其他女性都没有能够留下后裔。分子生物学家用什么方法证明了这些“夏娃姐妹”的线粒体 DNA 与之相同，又用什么方法证明了夏娃时代的非欧亚人的线粒体 DNA 存在不同类型？既然现代人的线粒体 DNA 只有一种类型，那么分子生物学又能凭什么判断还有别的种类的线粒体 DNA 呢？

因此，夏娃理论的成立必须要有第一预设——当时非欧亚智人的线粒体 DNA 各不相同。群体遗传学家认为：如果一个物种的各个群体彼此保持地理上的接触，由突变引起的遗传变化可以借助于杂交而传播到整个地区，其结果是虽然这个物种的遗传构成得到了改变，但整个物种在遗传上保持一致；与之相反，如果一个物种的各个群体由于自然巨变导致地理上的隔绝，那么起始于一个群体的遗传变化就不可能传播到其他群体，遗

传上彼此不同的结果最终导致不同亚种或不同的物种的形成[1]。由于猿人分化较早和远距离隔绝，可以推定夏娃时代的非欧亚人类的线粒体 DNA 是各不相同的，他们是人类进化过程中的不同亚种。

既然夏娃理论的倡导者依据地理隔离形成亚种的生物学公理推导出第一预设"夏娃时代人类线粒体的 DNA 各不相同"，接着就必须具有第二预设——原人女性所生育的子女被夏娃人种赶尽杀绝而绝不可能延续至今。在换代过程中，土著原人是被鲜血淋漓的杀戮，还是自生自灭的消亡？根据人类学家朱布罗(E. Zubrow)的群体影响替代模式，一个强势群体只要有 2% 的优势就可以导致1000年内另一个弱势群体的消亡。人们很容易理解通过暴力优势毁灭一个群体，但很难理解由于经济弱势的一个群体的自然衰亡。尽管朱布罗并不能明确地显示出人口统计学上的现代人祖先和欧洲尼人的取代方式，但是确实揭出暴力不是造成种族取代的唯一模式。在西欧的 3.5 万年前，尼人与新人共存期间长达 2～1 万年，他们的换代似乎符合朱布罗模式[2]。依据阿德里(R. Ardrey)《非洲的创世纪》中的著名开场白："人类不是生来就清白无辜的"，仍然可以设想：群体影响替代模式并非完全排除暴力机制，也许种族竞争的替代过程是多种机制在起作用。

为了使这种换代过程减弱一些残忍的兽性而增强一些温和的人性，还是可以另辟径蹊：既然人类的线粒体 DNA 都来自母系遗传，父系不可能进入受精卵，那么夏娃母系后裔在向世界流播的过程中，土著父系与之结合也就不可能在线粒体 DNA 中

① [美]理查德·利基《人类的起源》，69 页。

② [美]理查德·利基《人类的起源》，75 页。

留下任何痕迹。因此夏娃理论虽然排除了现代人母系多源论，然而并未排除父系来源的复杂性。换而言之，即使分子生物学确证现代人来源于同一母系，但是并没有否定现代人可能来源于不同父系。除了与夏娃同系的"亚当"，与夏娃不同母系的"亚当"是不是有可能是当地原人呢？或许，正是这些不同母系的人种之间的杂交才产生了更为智慧的现代人。可以把这一过程设想得复杂一些：由于土著母系的后代经过多次反复杂交逐渐消融在夏娃母系基因之中，最终才形成了现代人线粒体基因的一致性。因此是否可以提出"同系夏娃·不同亚当"或"反复杂交·逐渐消融"的假说呢？尽管威尔逊说他们的结论在统计学上是无可否定的，可是近来发现了一些统计上的问题，其结论被认为不像曾经宣称的那样肯定。"现代人类线粒体DNA基本相同性"是不完全归纳法的结论，如何解释0.32%的平均差异率呢？这些差异率是否是"反复杂交、逐渐消融"的孑遗呢？最后，在夏娃理论者那里还存在第三预设——即使新人移民与原人土著之间的杂交存在，其程度也是无穷小，即使发现不同类型的线粒体基因，对现代人的群体基因贡献也不大。

对于东亚大陆现代人的来源，中国考古人类学家在并不排除附带与境外人种杂交的同时坚持主体区域进化论，但是迄今为止东亚大陆一直没有发现距今10～3万年之间的新人化石。尽管人类学家坚持"没有发现不等于不存在"，合乎逻辑的说法是"没有发现无所谓是否存在"，但根据中国分子生物学家在东亚和东南亚采集的1500～2000份样本的基因研究表明，没有一例是由当地原人进化而来的后裔。[1] 由此假定北京猿人——山顶洞人的后代在7.5～1.5万年之间的冰川时期灭绝了，现代东

1　寿蓓蓓《中国人在哪里长大成人?》，《南方周末》1999年9月24日。

亚人来源于非洲新人。这种假说隐含着一个新的群体替代模式——由于自然灾变导致原人的灭绝而实现人种换代。

换代论和区域论，与其说是欧洲人类学家和美国人类学家的分歧，不如说是生物学家与考古学家的矛盾。考古学家不相信分子学的基因归纳，认为用基因分析的结果推测人类进化的过程是间接的；生物学家不赞成考古学的解剖比较，认为区域进化是把考古实物综合起来讲故事。针对"欧洲可能是换代的，但不能认为亚洲也发生了同样的事。中国人的进化是以连续进化为主，附带与境外人类的杂交"的观点，必须提醒：10 万年前并没有欧亚非各洲的人文地理分界，更没有民族和国家意识，在现代人的起源研究中务必抛弃潜在的"国族中心情结"。从达尔文开始，在人类起源的问题上，欧洲学者相继抛弃了"欧洲中心主义"，最近又提出白种人由黑种人在 3 万年前变化而来的假说。尽管为了进一步探索，可以不放弃换代论，然而无法改变这一事实——已知的最早人类化石都出自东非。就方法论而言，与考古人类学采用的是结论具有或然性的类比法不同，生物分子学采用的是结论具有或然性的归纳法。不同研究的印证才能增加结论的可信度，最终结果应是生物学与考古学各自独立研究的殊途同归。依据夏娃理论以及相关条件，现代语言的起源是单源的，有声胚胎语的形成时间是在距今 10 万年之际。

## 四、人类思维交际符号演变的连续统

根据儿童语言的发生学研究和聋哑人的交际方式等可以推知，在漫长的旧石器时代，声道没有成熟的人类所使用的是"非语音符号手段"。起初，从动物性信息交换方式中形成由面部表情、实物指示和身体动作等构成的喊叫伴随性原始体态语，然后在其基础上形成较为专门化的发音伴随性原始手势语，与体态

语和手势语阶段相应的是实物直观思维和手势符号思维。随着声道的逐步成熟,形成与体态语、手势语相伴随的表义音符增多性有声胚胎语,原始语言思维并始出现。在胚胎语进一步演变成为有声分节语言的同时,体态语和手势语也就相应蜕变为与语言交流相伴随的副语言,即现代意义上的说话时的表情体态。与副语言伴随性有声分节语言相应的是现代语言思维。距今3.5万年之际的新人文化,应当是建立在集体狩猎经济基础上的中石器文化和具有群体交流性的有声分节语言文化。

原始"非语音符号手段"和成熟"有声分节语言"之间,有一个漫长的过渡时期。随着大脑的日益发达和语言器官的最终成熟,"非语音符号手段"渐渐为次而"有声语言"日趋为主。咽腔的日益成熟、气流的自如控制使音节的离析成为可能,有声分节语言最终成为人类交际中的主导,在实物直观感受、手势符号感受和原始语言感受的基础上最终演进出现代语言思维,由此形成人类思维交际符号演变的连续统。

| 人科形成 | 能人时期 | 早期智人时期 | 体质新人时期 | 心智新人时期 |
| --- | --- | --- | --- | --- |
| 动物性交换方式 | 喊叫伴随性原始体态语 | 发音伴随性原始手势语 | 表义音符增多性有声胚胎语 | 副语言伴随性有声分节语 |
| 动物性感受(混沌意识) | 实物直观感受(形象思维萌芽) | 手势符号感受(抽象思维萌芽) | 原始语言感知(语言思维萌芽) | 现代语言感知(语言思维成熟) |

这一演变连续统的总趋势是:1.非语音符号的由强渐弱和语音符号的由弱渐强,语音符号最终覆盖非语音符号的同时,非语音符号变成副语言。人类意识形态从混沌性、经由实物直观性、动作意会性和原始语言性,演变成现代语言性。2.尽管如

此，都是作为个体的现代人的精神世界中仍然包含着不同层次的动物性直接感受（本能反馈、潜意识）、实物直观感受（形象思维）、动作符号意会（动作思维）、混沌语言思维和清晰语言思维，并非人类精神世界都可能或都必须以语言方式表现出来。3.语言思维成为现代人的本质特征（以学者为典型），灵感来自本能反馈或混沌意识（以巫师为典型），形象思维来自实物直观感受（以画家为典型），动作思维来自动作符号意会（以舞蹈家为典型），而音乐则来源于情感的原始声音宣泄方式。

## 五、现代语言造就了现代人类和现代文化

现代人类的语言起源和成熟，与现代人类起源的夏娃理论、现代人类的完全直立、现代人类的神经机制成熟、现代人类的男性体质成熟、现代欧洲祖先的来源、现代东亚祖先的来源、现代人类基因的生态变异以及现代人类的心智成熟密切相关。

### （一）12万年前：现代人类的女性体质成熟

美国威尔逊小组的科学家通过不同种族胎盘的线粒体DNA遗传研究，推定现代人类的线粒体DNA最终都是一个共同女祖的遗传。支撑夏娃理论的考古学证据，是10万年前的非洲新人化石和9～7万年前晚期旧石器文化。现代人类起源的夏娃理论也就蕴涵了现代语言起源的远古同源论，有声胚胎语的形成时间可能是在距今12万年之际。

### （二）10万年前：现代人类的完全直立及咽腔成熟

根据英国与荷兰科学家的研究，通过检查提供平衡感的内耳机制，发现现代人类10万年前的祖先才是第一种完全直立步行的灵长类动物。完全直立型姿态，口腔与喉部之间才有可能形成咽腔，现代人类才可能具备元音或音节的发音能力，为有声分节语的形成提供了生理条件。

(三)10万年后:现代人类的神经系统成熟

神经生理学家史密斯认为,一直到10万年前,人类祖先的神经生理状况和别的动物没有太大差别,然后一系列的前语言在某一方面的预先适应,与大脑进化到关键体积的相一致导致了语言的突然发明。现代人类的神经生理状况的成熟,表明符号化认知能力的成熟。

(四)9万年前:现代人类进入中东地区

在非欧亚通道的中东地区,发现了距今大约9万年左右的卡夫寨洞穴新人化石,表明现代人类走出非洲,开始向欧亚大陆迁徙。一方面,中东地区成为现代人类迁徙的中转站,另一方面,中东地区后来成为新石器革命的发源地、最早的农业定居点和文字创制的最早地区。

(五)6万年前:现代人类的男性体质成熟

根据生物分子学的研究,现代人类的男性体质成熟比女性要晚5.7万年,即在大约6万年前。这一男性进化稍后的过程,在人类胎儿的发育过程中有所反映。英国的一项研究表明,男婴出生时发育程度平均比初生女婴慢4—6周。男性体质的成熟,标志着现代人类体质的最终成熟。据此推定,现代语言在5万年左右趋于基本成熟

(六)4.5万年前:现代欧洲人的祖先进入欧洲

英国科学家根据线粒体DNA特征研究的分组结果,推测现代欧洲人的99%是由在4.5万年前迁入欧洲的7位女性祖先繁衍而来。按照线粒体DNA的特征划分,现在的非洲人分属三个"家族",其中一个"家族"与这7位欧洲女性祖先的血缘关系密切。

另外一些科学家在欧洲的25个不同地点,对1007位男性的Y染色体进行检测,发现欧洲现时人口可以归纳为10个"基

因部族",而其中的80%大约4万年前来自同一男性祖先。在距今2.4万年前的最后一次冰河时期开始时,这些早期欧洲人已经散居到西班牙、巴尔干半岛和乌克兰一带。这项研究显示,4万年前抵达欧洲的一个原始捕猎者,他的子孙逐步繁衍并且不断迁徙而支配了欧洲大陆。

考古研究表明,在西欧的3.5万年前,尼人与新人共存期间长达2~1万年。1997年,德国慕尼黑大学动物研究所的科学家宣称,他们对欧洲原人——尼人化石中的基因残存物的分析结果支持了非洲人迁徙到欧洲,取代了当地古人类的这一假说。同时,欧洲科学家提出了白种人由黑种人在3万年前变化而来的假说。

(七)4万年前:现代东亚人的祖先源于非洲

由中国国家自然科学基金资助,中国的一些研究机构和美国得克萨斯大学遗传学研究中心合作,在探讨中国种族群落遗传联系的同时对东亚人口的起源进行了研究。利用以30个遗传标志为对象的基因分析,获得了中国28个人口群落的遗传详情,然后再与全球其他11个人口群落进行比较。1998年9月29日,发表于美国国家科学院学报上的成果指出:东亚人口的遗传标记大都相同,表明他们起源于共同祖先,东亚人口的大多数源于非洲新人。根据东亚大陆一直没有发现距今10~3万年之间的新人化石,根据东亚东南亚采集的2000份样本的基因研究而无一例外,可以推定山顶洞人的后代在7.5~1.5万年之间的冰川时期灭绝了。

复旦大学生命科学院科研人员及其国际合作研究者在美国《科学》(2001年5月11日)上发表研究结果,通过对涵盖东亚、东南亚、大洋州、西伯利亚和中亚的163个人群的12,127位男性的Y染色体进行检测,在所有的样本上都发现了一个突变位点

M168G。这个突变位点大约在不早于7.9万年前产生于非洲，是一部分非洲人特有的遗传标志。之所以选择Y染色体，是因为这种性染色体在代代相传中大体不可能“重组”，总是稳定而全面地记录遗传信息。这是目前支持“东亚人非洲起源说”的最强有力的证据。

从世界人类演变史来看，100万年前世界性移民浪潮中，原人的祖先从非洲迁到中东、欧洲和东亚进化为当地古人。而基因分析的结果证明，最近一次世界性移民浪潮中的新人才是现代人类的直接祖先。目前的基因证据并不支持黄种人有独立起源的可能，源于非洲的现代人类在4万年前到达东亚，逐步取代了当地的原居古人。

(八)人口迁徙留下现代人类基因的生态变异

美国破译人类基因密码的塞莱拉公司与美国国家卫生总署的科学家宣布，根据基因草图，他们一致认为：种族是一个社会概念，而不是科学(生物学)概念，基因图谱并未显示种族之间有何差异。我们都是10万年前从非洲少数原始部落迁徙后进化而来。区别不同种族的常用特征，比如肤色、目色和鼻宽，是由极少数(0.01%)的外表基因所控制的。这些基因变异是为了适应环境所带来的巨大压力，而人的智力、艺术天赋和社交能力却是由数千个基因决定的。然而有人口遗传学家认为，种族的分类对于研究人类的起源和迁徙过程具有很大帮助[①]。世界上不同种族的人与人之间，99%以上的基因密码是相同的，现代人类的起源和进化是在相同基因的基础上完成的。根据不同地区的男性基因研究，发觉生态环境的改变可能引起某些基因的变异，

① 《科学家研究发现，基因证明人类无种族之分》，《扬子晚报》2000年9月12日

依据这些基因变异可以推定几万年来现代人类的迁徙时间和路线。

既然达尔文提出种群形成的序列将为各种不同语言提供最好的分类，那么根据现代人类起源的夏娃理论以及考古学发现，就可以尝试勾画现代人类的世界性迁徙历程以推导语言流播的影子。大约12万年前，东非某一聚落的黑妈妈由于基因突变有幸成为现代人类的“夏娃”。夏娃后代在10万年前开始迁徙。部分子孙滞留在非洲辗转（尼格罗种群的祖先），部分子孙向北走出非洲热带大草原的“伊甸园”，距今9万年左右来到中东，创造了中东新人文化。部分子孙滞留在中东（高加索种群中东型的祖先），部分在距今5万年左右进入东欧。在与欧洲尼人并存的同时，一支向西转徙，在3万年前出现在西欧（高加索种群欧罗巴型的祖先），一支向东迁徙到达东北亚（高加索种群乌拉尔型的祖先）。中东的一部分支向东，距今5万年左右经伊朗高原先后进入南亚印度次大陆（达罗毗荼种群的祖先）和东亚大陆。南亚的一部分沿孟加拉海湾北岸进入东南亚（蒙古利亚种群南亚型的祖先），有的继续东迁到南太平洋群岛（一连串的岛屿可能是陆桥），距今3万年左右向南直达大洋洲（大洋洲种群的祖先）。居住在黄河流域和北域的草原地带（蒙古利亚种群东亚型和北亚型的祖先），其中一部分继续北进，距今2万年左右到达北极（蒙古利亚种群北极型的祖先），又经白令海峡陆桥进入美洲（印地安种群的祖先）。在这一漫长的迁徙中，夏娃的后裔与欧亚当地进化起来的智人，不可避免地展开冲突、征服、趋同等方式的持久竞争，极其缓慢进行着的繁殖混杂使“夏娃基因”得以扩散，以大西洋和太平洋两大原始种群为过渡，在旧石器时代晚期最终形成地理生态种群。这一大致勾画出的现代人类的世界性迁徙历程，也就稀疏地投影出现代语言的起源、演变和流

播，关于现代人类语言的起源和新石器时代以前的语言流播，其研究大致也只能如此。

人类语言的历史并不那么漫长，有声胚胎语的出现不会早于10万年前，成熟的有声分节语言的历史可能仅仅只有4万年历史。原始劳动的历史、混沌意识、直观思维、动作性抽象思维的历史，比有声分节语言思维的历史要长得多。大约从1万年前开始，世界各地的狩猎——采集群体独立发展出的农业技术是文化技术进步的结果，而不是生物进化的结果。再追溯到这段社会和经济转变时期以前，会发现冰河时代的欧非的"狩猎——魔法"绘画和雕刻，这些心智产品显示出像我们现代人一样的精神世界。可是再往前追溯到大约3.5万年前，这些标志逐渐变弱而终于熄灭。人类学家相信，大约3.5万年前的考古记录中突然出现的艺术表现力和精巧技术性，是现代人进化的一种清晰信号。只有到这一阶段，当旧石器时代晚期文化突然出现在历史舞台上，创新的表现力和随心所欲施加的规则才变得普遍，开始用打下的石叶制作各种精细的工具。艾萨克（G. Isaac）认为，磨制工具必须施加一定规则，规则的施加是意志的执着，这是一种要求复杂口语才能够实现的行为方式。没有语言便不可能将规则施加其所制作的对象上[1]。有目的的集体狩猎、巫术魔法也都需要这一方便而神奇的符号系统。只有到了距今3.5万年之际，体质现代人才终于完成了到心智现代人的突变。

人类学家奥克利（K. Okelly）认为，现代人类行为的花朵就是与现代化语言的首次出现相联系的，难以相信，人类这个物种在其他方面具有完全现代特征而语言却不是现代的，现代人的

---

[1] ［美］理查德·利基《人类的起源》，102页。

出现是由于语言完善到我们今天经历的水平而引起一连串的连锁反应的结果。语言的成熟被确定为人性出现的过程中达到顶点的事件[1]。所谓"人性"就是语言性，所谓"人的世界"就是语言符号的世界，正是语言的出现才使世界变得如此清晰。也许就是人类历史上的第一次信息革命的成果——有声分节语言的成熟，才使现代人获得了有效的信息传播工具，掌握了开启文明殿堂的"芝麻开门"。

## 第二节 超越谱系树模式的语言关系类型学

尽管一些考古学家并不赞同依据基因检测提出夏娃理论假说，但大多数人类学家权衡各种依据，倾向于认为现代人起源于撒哈拉以南的东非。除非生物分子学被彻底推翻，否则现代人起源的总体状况不能改变。许多人类学家之所以赞成语言是晚近时期迅速出现的一种符号系统，正是因为注意到旧石器晚期革命中人类行为的急剧变化。如果坚持人类进化的区域论，那么必然推定语言起源的多源论，并以之解释现在所见的语言范型的多样性和复杂性，但是依据新古人类的换代论，则必然认定现代人类语言起源的同源论，并以不同群体的语言演变差异性来解释人类语言范型的多样性和复杂性。

### 一、语言关系类型历史演变的差异性

关于远古语言的存在状态，俄国民族学家 С·Л·托尔斯托

[1] 美 理查德·利基《人类的起源》，62 页。

夫(С. Л. Толстов)曾经提出过一个原始语言连锁扩散性假说[1]。洪荒之初的人类可能说多种语言,在民族毗邻地域逐一接连发生变化,形成语言连锁扩散网,某些地区不久之前还保留着的这一古代语言分布状况的痕迹。澳大利亚原有几百种土著语言,然而在它们之间很难找出明确的界限;在新几内亚的巴布亚人中,几乎每个村庄都有自己特殊的语言。相邻部落的语言虽有差别但能通话,可是在相距较远的部落之间的语言差别大到难以交流。这种在一片区域之中,发音和语法都相去甚远的不同语言,形成了一个原始语言连锁扩散网络的甬道。原始语言连锁扩散性假说可以成立,只是要给予一个有助于现代语言研究的大致存在时间——距今1万年或旧石器时期晚期,并且称之为原始语网。

从理论上来说,人种的基因扩散和语言的结构基因扩散应当是一致的,这两者都受控于早期的自然生态。根据现代人起源的同源论假说,可以相应假设现代语言起源的同源论,其扩散过程大致如下:1.距今10~6万年之间,在夏娃种群出现以后,随着人类完全直立过程中的咽腔的日益成熟和神经生理机制的日益发达,人类现代语言的原始语源萌生。2.距今6~4万年之间,随着新人的世界性迁徙,以乌拉尔山脉—乌拉尔河—里海—卡拉库姆沙漠—伊朗高原上的沙漠和山脉为界形成大西洋原始人种和太平洋原始人种。具有同一源头的现代语言向各地缓慢流衍,形成血缘型原始语流,旧石器时代晚期的现代有声分节语言日臻成熟。3.距今3~2万年之间,与环境影响所形成的尼格罗、高加索、蒙古利亚、大洋洲和印地安等五种生态人种相对应,

1 [俄]尼·切博克萨罗夫、伊·切博克萨罗娃《民族·种族·文化》(赵俊智等译),东方出版社1989年。

存在旧石器晚期五种生态地理人种型原始语圈:以撒哈拉沙漠为界分化为南部尼格罗语圈和北部高加索语圈,以印度洋—太平洋相连海域为界分化为大陆蒙古利亚语圈和大洋澳斯特列语圈,在美洲形成印地安语圈。4. 距今2～1万年之间,进入新旧石器交替时期。不同的生态环境进一步制约着经济生产方式和语言文化类型的发展,对自然世界和精神世界的分类认知导致了语言范畴的出现,新旧石器交替时期的原始文明中心逐步萌生。强势文化扩散传播的不平衡性表现为相邻语言的相似性和远隔语言的相异性,不同的血缘聚落语趋于不同程度的相连地域同化,形成地缘型原始语网。5. 在距今大约1万年以后,欧亚非的一些地区先后进入新石器—金属时期,随着社会分工的出现,从原始狩猎方式中形成游牧经济,从原始采集方式中形成农业经济,产品的多余萌发早期的商品交换,自源文字开始出现。在这种经济文化背景下,世界语言关系呈现出经济文化圈型差异格局:(1)牧畜民型的远距离迁徙突破原始语网,语言的分离和独立发展形成分化型语系;(2)农耕民型的文化扩散突破原始语网,语言的推移和互动接触形成同化型语区;(3)渔猎·采集民型(可能存在原始农牧萌芽)保留着连锁型语网;(4)海上游牧民型的迁徙不定和交往闭塞形成离散型语珠。

根据以上线索,提出"现代人语言关系类型历史演变的连续统假说":原始语源——原始语流——原始语圈——原始语网——"语系·语区·语网·语珠"四分格局。

| 10～6 万年 | 6～4 万年 | 4～2 万年 | 2～1 万年 | 1 万年开始 |
|---|---|---|---|---|
| 同源型 | 血缘型 | 生态人种型 | 地缘型 | 新石器—金属时期经济文化圈型 |
| 原始语源 | 原始语流 | 原始语圈 | 原始语网 | 牧畜民型分化型语系<br>农耕民型同化型语区 |
| 血缘聚落语 | | | 渔猎采集民型连锁型语网<br>海上游牧民型离散型语珠 | |

在西方学者看来，语系形成的时间对于民族起源的研究至关重要。某些考古学家和民族学家认为，语系可能形成于旧石器时代末期，距今已有1.3～0.7万年。在这一时期的人类迁徙过程中，某些大型原始民族的语群可能已在广阔的区域中传播开来。以游牧经济方式为主的闪含语言、印欧语言、乌拉尔语言和阿尔泰语言的分化可能如此，但是如果认识到世界语言发展的不平衡性，就不应当机械套用亲缘分化型模式。以往的历史比较语言学，仅把语系作为归纳语言关系的唯一的最终上位概念，因此除了勉强归纳出一些"语系"(其中不少是非语系)，在历史比较语言学已经建立200年后的今天，一些地区的语言或者划分太繁(例如印地安语言)，或者仍难纳入语系范畴(比如大洋洲原居民语言)而被称为系属不明。显而易见，西方语言学家引以为骄傲的谱系树理论是建立在印欧语料之上的一种语言关系类型。作为人类语言关系类型的第一种模式，在急于展开研究的同时而对一些基本理论至今没有深入探讨。因此，在促进世界语言关系研究的同时，也带来了不可避免的负面影响。

之所以以往的语言同源研究治丝益棼，可能是因为以下三

个主要原因。第一,没有考虑语言的同源年代性。语言的同源至少可以分为:1.旧石器时期的现代语言起源的同源;2.新石器—金属时期或文明中心形成以后的现存语言起源的同源。为了进一步明确语言的同源性研究,不妨分解为两个命题:现代语言的起源性研究和现存语种的发生学研究。前者可以称之为现代语言的同源性研究;后者可以称之为语种的发生学研究。语系研究的本质上也是语种的发生学研究,只不过这是在历史上曾经分化为不同子语的、某一种十分古老的原始语种。源头年代的相对性决定了语言起源的相对性,研究语言的同源务必明确:是什么时期和什么性质的语言同源性。第二,没有明确语言的同源性包括在"语言的相似性"这样一个更大的命题中。语言的相似性可划分为同源相似和借源相似。语言的同源和借源具有相对性,历史比较语言学的预设——语系的同源性——只能是指新石器—金属时期的语言同源性。以金属时期为标准所确定的同源,有可能在新石器时期是借源;以文字记载为标准所确定的同源,有可能在前文字时期是借源。为什么历史比较语言学家在同源词和借源词上争执不休?不得不笼统地称之为"关系词"?最根本的原因就在于缺乏年代性标准和同源相对性观点,而没有年代标准就不可能考虑同源和借源的相对性。第三,没有认识到世界语言关系类型的复杂性。如果以新石器—金属时期为上限,除了处于语系模式中的子语具有同源性以外,处于其他关系模式之中的语种之间可能不具有同源性,也就是说这些语言没有新石器—金属时期的语系式祖语。如果这些语言之间具有某种"同源性",那么可能是扎根于旧石器晚期的同源性。因此,我提出面对世界语言的"语言关系类型学",并且明确"语系"模式不过是"语言关系类型学"中的一种模式,从而为世界上不同语言之间的不同关系的全面解释拓宽了道路。

## 二、世界语言关系类型学的四分法

新的狩猎与采集场所的寻求,在促使新人展开世界性迁徙的同时现代语言也就逐步流播到世界各地。在旧石器时代末期,尼格罗语圈可能已经分化为科依桑—班图—苏丹等原始语网,高加索语圈可能已经分化为闪含—欧罗巴—高加索—卡特维尔—乌拉尔等原始语网,蒙古利亚语圈可能已经分化为南岛—南亚—藏缅—阿尔泰—北极等原始语网,印地安语圈和澳斯特列语圈也可能已经分别演化为原始语网状态。新石器革命的成功,导致古老的狩猎——采集经济方式在河流平原地区产生原始农耕经济,在丘陵草原地带形成原始游牧经济。距今1万年左右,两河流域出现了人类最早的农业性定居。前4000年到公元前后,随着金属时代的到来和自源文字的形成,随着商业的出现和早期国家的萌芽,人类游牧——农耕——通商文化的早期中心同时形成,并进一步组成金属时期语言文化圈。欧亚非交汇处出现古埃及—苏美尔(巴比伦等)—米诺斯(古希腊等)等文明中心,欧洲呈现日尔曼—黑海—俄罗斯等文明中心,亚洲东部出现哈拉巴—夷越—华夏—西域—阿尔泰等文明中心,美洲大陆出现玛雅—印加等文明中心。这一切都在促使迈进文明门槛的诸多语言冲破原始语网,导致语言关系类型向新的多格局方向发展。

(一)语系的形成:新石器—金属时期以来最典型的语言文化的同源分化现象

1.闪含语系。在公元前4000年就存在的含米特古埃及语,公元前后消亡后演变为科普特语。以早期游牧—通商文明为特色的含米特语言由西亚而南下,在公元前1500～1000年进入北非。随着闪含人的通商活动,闪含语言广泛分布在亚非地域形

成闪含语系。

2.印欧语系。公元前3000年,欧洲东部出现最初可能是原始农业经济的原始印欧文明,尔后逐步发展为以饲养为主导的游牧经济。游牧经济需要更新牧场,前2500年原始印欧文明开始分裂[①]。向地中海方向南迁的成为希腊—罗曼语言的祖先,向北海—波罗的海方向北迁的成为日尔曼语言的祖先,向东北方向迁徙的成为斯拉夫语言的祖先,向里海方向迁徙的成为伊朗—印度语言的祖先,由此形成地域广袤的印欧语系。

公元前1世纪,日尔曼语分化为斯堪的纳维亚半岛的北支、易北河与奥得河之间的西支和奥得河以东的东支。5世纪,日尔曼西支的盎格鲁、撒克逊、朱特人越过北海,进人不列颠形成古英语。9世纪,日尔曼北支的一部分西传成为古冰岛语。公元前1000年,罗曼人来到巴尔干半岛建立了罗马文明。5世纪,随着罗马帝国的崩溃,地区拉丁语成为独立语言。公元前10世纪,凯尔特语通行在德国南部阿尔卑斯山以北地区。前5世纪,凯尔特人西迁到英伦三岛。5世纪盎格鲁人人侵以后,凯尔特人后退到西北部形成威尔士语和爱尔兰语。部分凯尔特人在6世纪越海迁至法国西北部,成为布列塔尼语。在维斯杜拉河语第聂伯河之间的斯拉夫语言,公元前7世纪以后开始迁徙。西支迁至易北河与奥得河流域,南支迁至巴尔干半岛,东支进人俄国。波罗的海语言在前2000年进人俄国西部,由莫斯科附近的奥卡河一带北迁波罗的海。6世纪,波罗的海人为斯拉夫人所迫而西迁。到10世纪,立陶宛和拉托维亚语言已经不能互通,其他几种波罗的海语言在几百年前就消亡了。原始印欧语

① 参见[美]王士元编《语言与人类交际·印欧语》(游汝杰等译),广西教育出版社1987年。

的一支向东南到达里海，前 20 世纪进入伊朗和阿富汗，成为伊朗语言的祖先；前 15 世纪进入印度的“雅利安语言”，成为婆罗米语言和梵语的祖先[①]。继续东迁的进入西域和敦煌，成为吐火罗语的祖先。

3. 乌拉尔语系。公元前 4000 年，乌拉尔人居住在俄国中部，乌戈尔（Ugric）的词源来自古俄语“西伯利亚”的名称乌戈拉（Ugra）。前 3000 年，部分乌拉尔人开始迁移。迁向西北的乌戈拉语，在 1 世纪来到芬兰形成爱沙尼亚—芬兰语。说非芬兰语的原居民拉普人被迫迁入北极。迁往北方的乌戈尔语形成科米语。由北向东迁移的乌戈尔语，在鄂毕河—伊尔蒂什河流域成为奥斯恰克语，在叶尼塞河流域的成为萨莫耶德语。萨莫耶德人又继续东迁，长期居住在西西伯利亚泰加森林地带。公元初，部分乌戈拉人北移直到极北地区。公元前的最初几个世纪，部分乌戈拉人又折回西迁欧洲，9 世纪到达匈牙利的成为匈牙利语的祖先[②]。

4. 阿尔泰语系。在远古时代的东北亚山脉的森林地带形成阿尔泰人。随着游牧经济生活方式，西迁带里海一带，北迁到叶尼塞河流域，东迁到蒙古草原、黑龙江流域一带，形成阿尔泰语系。在公元前 5000 年就已经形成的东亚北方细石器考古文化系统，可能就是阿尔泰人创造的早期文化。据研究，华夏民族的始祖就是中国北方原始阿尔泰游牧民族的黄帝。古史传说中的黄帝与蚩尤之战、殷商时期中原民族与北方民族的战争，都是黄淮农耕文明和北方阿尔泰游牧文明之间的冲突交融。

---

① [美]肯尼思·卡兹纳《世界的语言》（黄长著、林书武译），北京出版社 1980 年。

② [美]肯尼思·卡兹纳《世界的语言》。

（二）语网的蔓延：滞留在旧石器—新石器早期的语言文化的连锁传播现象

1.原始达罗毗荼语网。公元前2000年，达罗毗荼语言就已经存在于印度的中北部，哈拉巴文明的创造者可能就是达罗毗荼人。雅利安人的入侵迫使达罗毗荼人南迁，在闭塞山区的达罗毗荼人及其语言残留下来了。达罗毗荼语网由于雅利安语言的冲决而残缺不全。

2.原始南亚—南岛语网。从印度次大陆一带的蒙达，到东南亚的孟高棉—马六甲—尼科巴诸语言，从东亚大陆东南部（朝鲜半岛—辽东半岛—山东半岛—江淮平原—东南丘陵为大陆系列）到诸海岛（日本—琉球—台湾—印尼—马来西亚—菲律宾为海岛系列），从马达加斯加直至南太平洋群岛诸语言，组成这一庞大的原始语网。据统计南岛语有700多种，南亚语有100多种。南亚—南岛语的故乡在东亚大陆的南部。从新石器早期开始，南岛居民开始了缓慢的南迁活动，先达到台湾和东南亚，然后一部分向西经印度次大陆直至非洲东部的马达加斯加，一部分向东到达南太平洋群岛直至圣诞岛。

3.原始巴布亚语网。16世纪前，巴布亚语就与迁来的南岛语部分交织。或建议用“印太诸语言”来包括这一地区的“非南岛诸语言”，在46万平方公里的土地上竟然集中了那么多的语言，在语言学上素有“未知领域”之称[1]。得到识别的语言已有650～850种，已证实有远亲关系的450种，归纳为76～105个小语系（sub-family），格林伯格将其划为7大语群。其余的200种大部分被划分为大小不等的语群，但有50种被认为是独立语

① 黄长著编著《各国语言手册》，重庆出版社2000年。

言,另外的则系属不清[1]。

4.原始大洋洲语网。在欧洲人进入大洋洲以前,原居民有500多个严格地守着各自"家园"的部落,也就存在着500多种不同的语言。1770年以后,随着被欧洲人的杀戮驱赶,人口从1788年的30万锐减到1960年的4万,语种急剧消亡,语网不复存在。据统计,现在仍说的土著语言还有253种,绝大多数濒临灭绝。较重要的语言,有通行于中部埃尔湖以北的阿兰达语(现存仅几百人)和中部最北地区的门金语。有人主张把它们划分为28个语系,可谓荒唐之极。有人主张看作一个联系松散的大语系或"澳大利亚大语群"(Australian Macro-Phylum)[2],其实就是语网。

5.原始印地安语网。现存和可考的印地安语言约有1000种,占世界现存语言的15%,被归纳为近110种语系(北美印地安语言70个语系,中美印地安语言20个语系,南美印地安语言24个语系),格林伯格将其归纳为10个。不是语系模式而是语网模式的印地安语言,与大洋洲语网一样,也早被欧洲人的迁入而变得支离破碎。

6.高加索语网。自古以来,处于黑海和里海、伊朗高原和俄罗斯平原之间的高加索山区一向被称之为"语言之山",现存语言约有50种。作为欧亚文明中心的边缘地区,高加索语言的结构类型代表着现存欧洲语言的最古老复综·粘着类型,北部语言比南部语言(格鲁吉亚语)更为古老。也许,欧罗巴语言和高加索语言在新旧石器交替时期同属"欧罗巴—高加索语网",由于原始印欧文明的崛起,原始印欧语从语网中分离出去,其语言类

---

① 《中国大百科全书·语言文字》,大百科全书出版社1988年。

② 参见黄长著编著《各国语言手册》,622页。

型转变为屈折语，并进一步分化为庞大的语系，而高加索语言成为处于横跨欧亚大陆的印欧语系汪洋中的惟一语网。

7. 科依—桑—班图—苏丹等原始语网。一些人主张分为科依—桑语系（15～50 种）、尼日尔—刚果语系（900 余种语言，班图诸语言有 150 种）、尼罗—撒哈拉语系。据统计，非洲现存语言有1900种，占世界现存语言的 32%。尼日利亚现存 410 种语言，乍得现存 210 种语言。苏丹、坦桑尼亚、埃塞俄比亚、中非等国家都有 100～160 种语言①。

（三）语珠的散落：语网的断裂后的语言文化的离散游移发展现象

虽然语珠的早先状态可能是语网中的语链 $a_1 - a_2 - a_3 - a_4 - a_5 - a_6$……，但是或者由于 $a_1$ 与 $a_2$ 之间的地域距离本来就很遥远致使语环松散，或者由于语种 $a_2$、$a_4$、$a_5$ 的远迁和相连文明的衰亡造成语环脱裂，以致于形成了 $a_1$……$a_3$……$a_6$……的语种零星分布状态，或 $a_1$……$a_6$……$a_3$……的语种错位状态。尽管 $a_1$……$a_3$……$a_6$ 稀散分布在一个相同的生态区域，由于环节语言的脱落，在现存语种之间已经很难或不可能找到语言的相似连接性。然而，在这些散落的珍珠之间也并非所有的相似性荡然无存，至少语言类型可能大致相同。

1. 北极圈语珠。居住在西伯利亚冻土带到阿拉斯加、格陵兰北极圈内外的一些语言，以往称之为“古亚细亚语言”。其祖先约于 5000 年（甚至更早）前从亚洲北部迁到白令海峡一带，以渔猎为生的这些沿海民族，其经济方式被称为“海上游牧”。追逐海洋猎场的漂泊不定，气候和地理的不便交通、长期缺乏文化

① 参见 Michael Krauss《危机中的世界语言学》（王朝晖译），《云南民族语文》2000 年第 4 期。

和语言的接触,北极圈语网的破裂则形成阿伊努(24,000人)—凯特语(1000人,据说是匈奴语的后裔)—尤卡吉尔语(800人)—尼夫克语(4000人)—科里亚克语(8000人)—楚科奇语(12,000人)—阿留申语(5500人)—爱斯基摩语(97,000人)的零星语珠[①]。

2. 苏美尔—米诺斯—巴斯克—亚美尼亚语珠假说。通常认为,前第5千纪分布于两河流域苏美尔人不知来自何方,前第2千纪分布于爱琴海上创造克里特文明的米诺斯人也是来历不明。这两种在公元前2000~1300年开始消亡、延续到公元前1世纪的语言,其结构都是粘着型。前2000年由小亚细亚移民带来的现存法国的粘着型巴斯克语,主要采用后缀、做格等语法手段。虽然现代亚美尼亚人是在前7世纪,由乌拉尔图人、斯基泰人(伊朗语族)结合而成的混合民族(现存人种属于巴尔干—高加索混合型),但前16~15世纪古代亚美尼亚人已经分布在亚美尼亚高原,现代亚美尼亚语实际是一种词汇印欧化,但结构仍保持着粘着型的混合语。根据语言结构的相同性(粘着语和做格),我提出"苏美尔—米诺斯—巴斯克—亚美尼亚语珠假说"(可能还包括埃兰、胡里特、伊特剌斯坎等古代语言)。如果这一假说得以证明,那么欧洲语结构类型的演变轨迹可能是:复综·粘着型(高加索语)—粘着型(苏美尔—米诺斯—巴斯克—亚美尼亚等)—屈折型(印欧语)。

(四)语区的推移:新石器—金属时期以来最典型的语言文化的异源同化现象

1. 东亚大陆语区。如果注意到东亚大陆以农耕经济方式为主,依据桥本的语言地理类型学,可以提出"地缘同化型语区"模

---

① 参见《民族词典》,上海辞书出版社1987年。

式。距今7000年左右,东亚大陆存在南亚—南岛语、藏缅语和阿尔泰语。在这三大语言交汇之处的中原,从前3000年到公元前后混合而成华夏汉语。以文明辐射为背景,在强大语言文明中心和周边语言文明的持久互动接触中形成东亚大陆语区。在这一互动过程中,以华夏汉语为中心可以合称为“华夏汉语的阿尔泰和或阿尔泰语的汉语化,与华夏汉语的南亚化或南亚语的汉语化”。形成这一状态的原因是:(1)农耕文明与游牧文明的冲突交融;(2)中原农耕文明的发达;(3)北方游牧文明的善于征战;(4)南方农耕文明的不断再南迁。语区的推移和同化是这一广大区域语言演变的总态势,但并不排除同时存在局部的分化。

2.次语区。所谓“次语区”是指欧洲语言学家在19世纪认为不可思议的畸形现象、后来被称为存在于大语系中间的“语言联盟”。最典型的是“巴尔干次语区”,虽然巴尔干半岛的这些语言(独立语族的希腊语和阿尔巴尼亚语;罗曼语族的罗马尼亚语;斯拉夫语族的保加利亚语以及部分涉及斯拉夫语族的塞尔维亚—克罗地亚语和阿尔泰语系的土尔其语),除了土耳其语以外都属于印欧语系的不同语族,但长期共处同一地域,拜占庭文化的凝聚力和各语言之间持久深人的接触渗透,致使这些语言面貌具有地域趋同性,形成从词汇到音位系统和句法形态的“巴尔干语言特征”。

语言关系类型学的四分法把语言关系类型确定为上位概念,下位概念是语系、语区、语网和语珠。1.语系的基本概念包括祖语(新石器—金属时期的语言源头)和子语(历史上或现存的语种)、近亲语(同源相似性密切的语种之间)和远亲语(同源相似性疏远的语种之间)。语系模式的层次是:语系(祖语)——语族(语支的源头)——语支(语种的源头)——语种(子语)。2.语区的基本概念包括核心语(新石器—金属时期文明中心的

语言)和周边语(周边文明中心的语言)、覆盖语(语言同化中覆盖底层语言的强势语言)和底层语(语言同化中所保留的弱势语言)。语区模式的层次是语区(没有新石器—金属时期的祖语或可能具有旧石器时期的遥远祖语)——语类(相似的语组)——语组(相似的语种)——语种(单一语言)。3.语网的基本概念包括环节语(处于同一语网中的某一语种)和连锁语(处于同一语网内的相连语种)、近邻语(处于同一语网内的相邻语言)和远离语(处于同一语网内的不相邻语言)。语网模式的层次是语网(旧石器时期的祖语遥远)——语链(相邻的语环)——语节(相邻的语种)——语种(单一语言)。4.语珠的基本概念包括语珠(由远距离的语种组成)和语种(单一语言)。如果采取语言关系类型学四分法,那么也许对从新石器—金属时期以来历史上及现存世界语言的关系状态可能会有新的认识。

## 三、语言结构类型历史演变的连续统

新石器—金属时期以来,随着表达精密的需要、语法范畴的消长、语言接触的影响,现代语言的结构类型发生演变。据现存的结构类型推断,最为古老的结构类型可能就是保留着混沌性思维烙印的复综语。这是一种词单位与句单位尚未分离、构句单位为"句词"的原始类型。就其结构而言,句词相当于其他结构类型中的形态复合词;就其功能而言,句词相当于其他结构类型中的句子单位。

以复综结构为出发点,语言结构类型的渐次演变大致如下:1.复综结构向分析结构演变。复综词逐步分解为语素结构(如中美印地安ABC三种语链复综结构的递减);复综结构的语音编插方式,如演变为词单位的语音屈折则形成屈折结构型;复综结构的语义音节化,如蜕变为词根和词缀相结合则形成粘着结

构型。2.屈折结构由内部屈折向内外屈折演变。内部屈折语由辅音构成根词干、以元音屈折为变化手段,内外屈折语在内部手段之外又出现了辅音性和音节性附加词尾。3.粘着结构由紧密粘着向松散粘着演变。紧密粘着型的词缀对词干依附性极强,松散粘着型的词缀与词干的结合趋于离散化。4.松散粘着型向孤立化结构演变。松散粘着型的词缀虚词化和词干实词化,有可能蜕变为以语序和语气为主要手段组合实词和虚词的孤立型,如历史上的汉语。5.屈折语由综合型向分析型的演变。内外部屈折结构随着内部屈折的消亡、外部词尾的失落和功能词的增多,屈折型演变为以语序和语气为主要手段组合词干和功能词的孤立型,未来英语的趋势可能如此。

人类语言的结构类型演变的总趋势是:随着语义表达的清晰化,语言结构呈现出语言单位的分析化、语法手段的附加化、语法成分的音节化和结合方式的孤立化的演变倾向。具体来说:1.语言单位的演变由复综性(复综型句词)到分析性(屈折型形态词——粘着型词根与词缀——孤立型实词和虚词);2.语法手段的演变由内部屈折法(闪含语的三辅音根词的元音屈折)到外部附加法(印欧语的辅音性和音节性词尾);3.语法成分的演变由音素性(音素交替——辅音性词尾)到音节性(屈折型音节性词尾——粘着型音节性词缀——孤立型音节性虚词);4.结合方式由附加法的粘着性(粘着型词缀)到语序法的孤立性(孤立型虚字)。

虽然从理论上可以假定纯粹结构类型的存在,但在现存语言中却找不到纯粹类型,可以看到的都是以某一类型为主或不同典型之间的过渡,由此形成了结构类型历史演变的连续统状态:

| | | | | |
|---|---|---|---|---|
| | 复综－屈折型 | 屈折型（内部屈折型） | 屈折 粘着型（内外屈折型） | |
| | 从复综性到语音屈折性 | 以内部音素屈折为主 | 从音素性屈折到音节性附加 | |
| 复综型 | 复综词·语音屈折（非洲弗拉尼语？） | 词干·语音屈折·词缀（闪含语言、格鲁吉亚语、波罗的海语言） | 词干·屈折·词尾·功能词（现代英语） | 孤立型 |
| 词单位与句单位的混沌性 | 复综－粘着型 | 粘着型（紧密粘着型） | 粘着－孤立型（松散粘着型） | 以语序法和虚字（功能词）法为主 |
| 复综词（中美印地安C类语言、爱斯基摩 阿留申语言、巴布亚语言） | 从复综性到外部音缀粘着性 | 以外部词缀粘着为主 | 从粘着性附加到离散性虚词 | 实词和虚词（汉语、壮侗－苗瑶语、高棉语、越南语）或词干和功能词（未来英语） |
| | 复综词·词干·音缀（北高加索语、北美印地安语、中美印地安A类语言） | 词根·词缀（乌拉尔语、达罗毗荼语、南岛语、阿尔泰语、南美印地安语、班图语） | 词根·词缀·虚词（亚美尼亚语、南亚蒙达语、藏缅语） | |

语言结构类型的历史演变过程及其结果非常复杂。1.确定某种语言的结构类型，只就主要特征而言。2.在后一类型的语

言中包含着不同历史层次残留的前类型孑遗(如藏语)。3.层累性、多向性的不同结构类型的接触,导致以一种类型为主的语言中包含着接触语言结构类型的不同特点(如原始华夏语)。4.越是混合的语言形态失落越多,导致由综合向分析(如英语),由形态向孤立的结构转型(如汉语)。5.孤立结构与屈折或粘着结构的深人接触,可能导致孤立语的再转型(如五屯话)。

在史前时期,一定的语言关系类型与一定的语言结构类型具有对应性,一定的语言关系类型和语言结构类型这二者与一定的经济方式具有相关性。直至15世纪,世界性的语言关系类型与语言结构类型的对应性还可以大致看出:

| 关系类型 / 结构类型 | 大洋洲、美洲、非洲采集渔猎文明为主导的语网 | 北极地区渔猎文明为主导的语珠 | 欧亚非游牧通商文明为主导的语系 | 东亚南洋农耕文明为主导的语区 |
| --- | --- | --- | --- | --- |
| 复综型 | 大洋洲-巴布亚语;印地安语;高加索语(与印欧语有远祖关系) | 古亚细亚语(与印地安语有远祖关系) | | |
| 屈折型 | | 闪含-印欧语言(地中海文化圈) | | |
| 粘着型 | 苏丹-班图-科依桑语(?) | | 乌拉尔-阿尔泰语言(处于欧亚之间) | 达罗毗荼-蒙达-南亚-南岛语-藏缅 |
| 孤立型 | | | | 壮侗-华夏苗瑶-越南语 |

随着不同文明之间的冲突同化日益加剧，语言的接触交融的范围也越来越大，地缘性的语言关系类型和认知性的语言结构类型之间的对应性逐步打破，人类语言卷入了从地域趋同化到世界趋同化的澎湃潮流。

## 四、人类语言文化的世界化进程

在地中海文化圈，两河流域的泥板尘封着人类最古老的苏美尔，定期泛滥的尼罗河水流淌着金字塔的碎影，克里特岛屿引起海洋的探索，腓尼基字母染就智慧的思想，巴尼尼语法宣扬信仰的圣歌，维纳斯女神从爱琴海浪花和奥林匹山峰中冉冉升起，古罗马狼嗷在阿尔卑斯山和博斯普鲁斯海峡之间回荡，耶稣的圣诞光大摩西的教义，阿拉伯世界向欧洲回流着古希腊的智慧，但丁吹奏着神曲驱散中世纪的阴霾，哥伦布的扬帆济海导致殖民主义的横行，蒸汽机的云蒸霞蔚呼唤着科技文明的日出。在东亚文化圈，文明的混合催生出华夏三代，黄河的水土哺育着盛世汉唐，西通丝绸东浮瀛洲，北至黑水南下红河，长城挡不住大漠的铁蹄，中原逐鹿导致士族的南迁，汴洛文明在江淮的山青水秀中再生，民族的冲突融合促使中华日益壮大。既然有文明的传播就有语言的接触，那么语言的交融从远古时期就开始了，其程度的深浅和时间的久暂导致语言演变的不同结果。

（一）语言的借用

语言接触过程中最自然不过的是借词，除了封闭的文明，世界上没有一种语言没有借词。持久的大量借词必然引起语言的变化，但只要作为独立国家或民族存在，即使同化程度再深仍然可以保持语言的独立性。尽管朝鲜语和日本语的词汇可能有70％来自历史上的汉语，但基本词汇和结构类型没有发生变化，因此仍然视为独立语言。作为语言的细胞的词，语音系统和语

法要素都包含其中。在借词过程中，语音和语法都可能发生结构关联性变化。受汉语影响，清代满语的音位系统就发生了很大变化；南方民族语言向汉语直接借用了许多虚词语法成分。如果主借和被借语言是两种不同类型，语法成分的持续借用可以导致结构类型的转变，如受藏语等语言影响从孤立转型为粘着的青海五屯话，受汉语影响从粘着转型为孤立的海南回辉话。

（二）语言的换用

说两种不同语言的居民长久共存于同一地区，会自然形成双语制。在一定情况下，双语制可能成为弱势语言被强势语言替代的中间过程，最终导致语言的换用。除了东北少数地方的满人保留着满语，关内满人语言的换用经历了满语制——满·汉双语制——汉语制的渐变过程。为了在中原长久统治的北魏王朝，提出与汉人联姻和"断诸胡语"的国策，导致了中原鲜卑语的主动消亡（据研究，清代西迁新疆的锡伯族即鲜卑后裔，锡伯语是延续至今的北留鲜卑语）。任何换用都不可能是纯粹的换用，换用者总会以不同方式在新换语言中打上原有母语的烙印，由此形成语言换用过程中的双重性：对于弱势语言的语主来说是语言的换用，对于强势语言的语种来说则是语主的部分变换，以致于形成换用语言底层和表层结构的双重性，而成为强势语言的一种方言（印度英语就是以印度人母语的发音方法，将英语改造而成的一种方言）。汉语的方言，绝大部分在历史上就是这样形成的，因此可以称之为对周边语言的"汉语化"不足。就不同语言之间而言，语言的借用和换用都是指语言间的同化，只是同化的程度不同。就某一语种本身来说，如果是强势语言，那么在同化别的语言的过程中仍然不可避免被弱势语言部分异化；如果是弱势语言，那么在被同化的过程中并非毫不异化强势语言。不管是强势语言还是弱势语言，都合乎"同化即异化"这一规律。

不同语言的同化与同一语言的分化是语言关系类型中的对立趋向，相邻不同语言之间的同化产生语区，同一语言的迁徙分化且分离发展形成语系。不但在语区形成中，“同化即异化”现象比比皆是，就是在语系分化过程中也在所难免。除非迁往无人区，否则远徙的某一语种在覆盖原住居民语言的过程中必然会受到异化，只是异化的程度与语区不尽相同。

（三）语种的消亡

除了自愿换用导致语言的自动消亡，民族文化的灭绝总是伴随着语言的被迫死亡。被阿卡德人打败的苏美尔人，其民族和语言的都隐没在历史的尘封之中。19 世纪中叶英国人来到澳大利亚南端的塔斯马尼亚岛，大肆捕杀原住居民，到 1876 年作为民族和语言的塔斯马尼亚，随着最后一位塔斯马尼亚妇女特鲁加妮妮的死亡而死亡了。经济弱势且人口稀少的种群的语言也难免自然消亡。人数在1500～4000之间的内婚制小种群，人类学上通常叫做德莫，这种滞留在文明社会之外的德莫语言，在历史上的消亡不知其数。

根据美国萨默语言学研究所的调查[1]，当今世界上只有一个人会讲的语言有 51 种，近 500 种语言的使用者不到 100 人，1500种语言的使用者不到1000人；3000多种语言的使用者不到1 万人；5000种语言的使用者不到 10 万人。全世界6000种语言的 96％的语种分布在 4％的人口之中。根据估计，其中一半左右将在 21 世纪消亡，即3000种语言将在1200个月中灭亡，这意味着每两周就有一种语言在世界上的某个地方消失。由此，人们发出“语种的消亡像物种的灭绝一样可怕”的哀叹！

---

1 戴维·克里斯特尔《语种消亡与物种灭绝一样可怕》，《参考消息》1999 年 11 月 10 日

(四)语言的混合

语言接触过程中最深入的是语言的混合。人类历史上的任何强大文明都是混合的文明。文明的混合必然伴随着语言的混合,语言的混合是历史上的自然过程,混合的结果则诞生新的语种。语言的混合可以分为两种类型:自然混合语和经济混合语。前者是指受同一地区语言的自然接触渗透,导致原语言不复完存而成为一种新语言;后者是指为了商贸活动出现的洋泾浜语,进一步的发展完善便成为克里奥尔语。

现代语言溯源于一,由于人口的流播、交通的阻隔、经济发展的不平衡等形成史前社会的上万种不同的语言。进入文明社会以来,从人类语言演变的总体趋势来说,语言分化只是局部或暂时的状况,语言运动最终导向语言的统一。从 15 世纪开始,欧洲科技文明冲出旧大陆向全世界洪溢,世界性的文化合流进程拉开帷幕。交通便捷、文化交融、通婚混血、经济合作和网络传播,使地球正在逐步变为"村庄"。一方面,世界性的经济、文化、科技、社会的发展需要统一的世界共同语或国际语;另一方面,语种的消亡正像物种的消亡一样不可阻挡。可以预见,国际语的形成并不是多种语言的杂凑,而是以一种语言为主体的混合。尽管国际语以何种语言为主体还不能断言,尽管国际语的最终形成可能还需要若干时间,但是从宏观趋势来看,语言的世界性合流不可避免,世界语言趋同化的进程仿佛重演着东亚地区周边语言"汉语化进程"的历史,随着历史上已经发生的文明接触交融和现当代正在进行的国际科技化,人类语言又逐步趋于统一。人类文明的发展由原始单一性走向民族差异性,再由民族差异性走向世界趋同性,人类语言的演变也必然由一源性到多样性,再由多样性到统一性:这就是人类文化和语言的宿命!

# 第三节　中国混合语的研究现状与理论探索[1]

## 一、混合语研究的历史回溯

正如房德里耶斯(J. Vendryes)所说:"使语言不受任何外来影响而不断发展的理想几乎从来也没有实现过。相反,相邻语言的影响在语言的发展中常常起重大的作用。这是因为语言的接触是历史的必然,而接触必然会引起渗透。"[2] 世界语言发展史表明,语言的演变往往与语言的相互影响密切相关。依据语言之间的影响程度,可以把语言接触归纳为三种类型:1.渗透。以借用方式造成的语言渗透。中国境内的语言接触大多属于双向渗透。2.换用。由借用对方的语言成分到直接使用对方的语言,经过双语制直到放弃本族语而换用对方语言。被换用语言以底层方式保留了本族语的孑遗,换用语言在融合对方的同时也发生了变异,从而出现了对被换用语言的人们来说是语言换用,而对换用语言来说则是语主局部嬗变的现象。3.混合。两种或多种语言之间强烈的相互影响,经历皮钦语和克里奥耳语两个阶段,形成了一种不同于来源语中任何一种的混合语。与大陆教材中通常的 pidgin 源于上海 business 的发音、creole 其义为混血儿的说法不同,据石定栩[3] 所引文献考证,pidgin 是广东发 business 的变音形成的;creole 源于拉丁语的 creare(创造),

① 本选题由李葆嘉提出,由张璇搜集资料写成部分初稿。

② 房德里耶斯《语言》(岑麒祥、叶蜚声译),310 页,商务印书馆 1992 年。

③ 石定栩《洋泾浜语及克里奥耳语研究的历史和现状》,《国外语言学》1995 年第 4 期。

葡萄牙人借作 criar(抚养孩子),在巴西葡萄牙语中演变成 crioulo(土著奴隶),后扩展到泛指在美洲出生的欧洲或非洲后裔,又引申为这些人的习俗和语言。这个词再借到英语里就成了 creole,慢慢地变得专指混合语。

石定栩在《洋泾浜语及克里奥耳语研究的历史和现状》中叙述,混合语的研究始于 19 世纪下半叶,第一个对混合语进行研究的是美国学者 Van Name(1869)。被誉为"克里奥耳语研究之父"的德国语言学家 Hugo Shuchardt(1880),第一次从理论上探讨了混合语形成和发展的机制。本世纪初房德里耶斯的《语言》(1920)和叶斯泊森(O. Jespersen)的《语言》(1922)都用一章的篇幅来讨论洋泾浜语和混合语。但是,许多欧洲人把洋泾浜语和克里奥耳语看作殖民地人民对欧洲语言的糟蹋,并以此来证明殖民地人民低能,无法学好宗主国的语言。不少欧美语言学家认为混合语是语言学习过程中的不良产物,没有学术研究的价值。殖民主义背景下的语言糟蹋论和历史比较语言学的谱系树模式眼光,致使混合语的研究长期以来没有任何地位。第二次世界大战中的美国,为了全球参战的需要,加强了对世界各种语言的研究和教学,在促进了语言学的发展的同时,也就促使混合语的研究进入一个新的阶段。美国 R. Hall(1944)是第一个正式研究中国洋泾浜英语的语言学家,又是为洋泾浜语正名而大声疾呼的第一人。他从理论上总结了洋泾浜语和克里奥耳语的生命周期,并且以混合语的研究结果来验证语言变化的理论,对历史语言学的方法论作出了贡献。1959 年出席牙买加西印度群岛大学的混合语研讨会的一些学者,自称克里奥耳学家,创立了克里奥耳学这个新词;1973 年夏威夷大学出版了《洋泾浜信使》;1989 年成立了国际洋泾浜语与克里奥耳语学会。混合语的研究已经成为国际语言学中不可忽视的重要部分。

20世纪50年代,中国语言学界以《马克思主义与语言学问题》为导向,像持有语言糟蹋论观点的许多欧洲人一样认为,该书认为两种语言的"融合"并不产生新的第三种语言,只能是征服式的一方"吃掉"一方。继而形成语言污染论在中国语言学界占据统治地位30余年的状况,"洋泾浜"被染上"殖民主义"的贬义色彩。近20年来,中国语言学界,主要是民族语言学界对中国境内的混合语言和语言混合进行了较为深入的研究。1981年,赵相如等发表《艾努语的数词——兼论艾努语的性质》,1982年,陈乃雄发表《五屯话初探》,由此引起中国语言学界对混合语的关注,有关论文不断发表,成为当代中国语言学和语言理论研究的一个新热点。

## 二、民族混合语研究述评

中国境内的混合语往往多产生在"西北……西南……云贵……广西……海南"这一语言文化走廊地带,而民族语言学家往往是混合语的发现者和研究者。这一语言文化走廊,不仅是汉文化与周边民族文化的毗邻地带,也是周边少数民族频繁接触自然杂居的地域,由此形成各具特色的混合语。1.汉语和阿尔泰语混合。青海汉语河州话具有汉语和阿尔泰语混合的性质①。2.汉语和藏缅语混合。青海同仁县一部分土族所操的五屯话。3.侗台诸语和汉语混合。广西融水县自称 $e^{55}$ 的壮族人,操一种五种语言混成的五色话②。4.南岛语和汉语混合。海南岛回族所说的回辉话,是占语和汉语官话的混合。5.白语和彝

① 马树钧《汉语河州话与阿尔泰语言》,《民族语文》1984年第2期

② 陈其光、张伟《五色话初探》,《语言研究》1988年第2期。

语混合。云南的卡卓语为白语、彝语混成的混合语[①]。6.阿尔泰语和印欧语混合。新疆南部艾努语是一种伊朗语在维吾尔语强烈影响下发生质变的混合语。

现对正在使用中的五屯话、卡卓话与回辉话这三种混合语的研究状况分别试加述评。

（一）五屯话：汉语和藏语的混合

五屯话是中国语言学界较早注意并研究得较为深入的混合语种。五屯话分布在青海省黄南藏族自治州同仁县，是当地"一部分土族居民"使用的一种长期受到藏语强烈影响的、以汉语为基础演变而成的混合语。五屯人的来历，陈乃雄起初认为最早来源于四川，后来有人指出五屯人来自江南。陈乃雄更改了原先说法，进一步证明，五屯人约在明初，最迟不晚于明朝万历年间，从南京一带迁来。受当地土族首领统治，经过几百年繁衍，今天已自称"土人"了[②]。

需要进一步明确的是，正像许多云南汉人自称他们的祖先是明代从南京迁去的一样，这些西南、西北江淮移民的祖籍"南京"并非限于如今的南京辖区，而是明代"江南省"的"大南京"，当包括以今南京辖区为中心的邻近苏皖江淮区域。由此推定，五屯人原说的江南汉语为明代江淮官话。依据通常看法，生产力与文化水平较低的民族在与生产力与文化水平较高的民族的接触中，后者的语言对前者的语言影响较大，其实这种说法并不确切。在特定的生态环境、特定的政治格局、特定的人口比例之下，影响力较大的语言必然是在当地政治上占统治地位、生产方式适应当地生态环境、人口占多数的民族的语言。南京或江南

① 戴庆厦等《云南蒙古族嘎卓语研究》,《语言研究》1987年第1期。

② 陈乃雄《五屯话音系》,《民族语文》1988年第3期。

的农耕生产力与文化水平虽然比青海地区的畜牧生产力与文化水平高得多，但是从江淮区域迁来的明代汉人，由于青海自然地理生态环境的原因，不得不放弃原有的农耕生产而改从畜牧生产。再加上政治格局中、处于被统治地位与人口比例占少数等因素，就不得不学习当地的通用语藏语，但又尽力保持母语汉语而不愿彻底换用藏语，以致原先的江淮官话久而久之就逐步演变成了后世的五屯话，江南人的后裔也就逐步自称"土人"了。如今，大多数五屯人是双语者，内部通行自己的五屯话，与藏民等交际时则用通用藏语。

根据陈乃雄的研究[1]，五屯话有 8 个舌面元音和 2 个舌尖元音。和汉语普通话一样，舌尖元音可以与舌尖前辅音和舌尖后辅音相拼。与普通话由于互补因而 2 个舌尖元音与[i]，通常处理为同一个音位不同，在五屯话中归纳为不同的音位。五屯话有 13 个复元音。有的源于汉语，如/ei/、/ia/、/ie/、/io/；有的源于藏语，如/yi/、/iu/。与元音受藏语影响较小相比，五屯话的辅音受藏语影响较大。五屯话有 33 个单辅音，很多不同于汉语普通话的辅音。如/z/只出现在藏语来源词里；而/ʎ/只出现在汉语来源词里。受藏语影响，五屯话存在 24 个复辅音，这些复辅音几乎全见于藏语来源词之中。只有/tq/和/ʎq/由汉语的/t/和/ʎ/演变而来，并且目前仍然尚不稳定，有时仍读作/t/和/ʎ/。声调在五屯话音系中的变异最为明显，可能由于多音节词的增多和受到没有声调的同仁藏语的强烈影响，五屯话基础来源语汉语的声调消失了，同时转化为一种词重音。如"买米"['mi me]（语序为 OV）、"卖米"[mi'me]，拆开来的[mi]、[me]并无轻重差异，组合时出现不同的相对轻重，这明显是声调消失后

[1] 陈乃雄《五屯话音系》

的孑遗或代偿。

五屯话中汉语源词占55%。藏语源词虽然只占20%,但是藏语源词已经渗透到了五屯话词汇的每一角落。在表达某些概念时,除了有汉语源词和藏语源词通用的情况,还有以汉语语素和藏语语素组合而成的汉藏合璧词。或在藏语词根的基础上接缀五屯话特有的词缀,如"出发"[lamalan]—[lamalan'la]。或在汉语词的基础上附接藏语词缀,如:"犯人"[lo 'ke wa ],其中的[lo 'ke ]源于汉语的"劳改",[wa]是藏语表人的词缀。虽然汉藏通用词与合璧词在五屯话词汇中仅占5%,但是异源通用词和合璧构词法却是五屯话在词汇和构词法方面的独有特征。以上这些语音和词汇的特点,都鲜明地表现了源于汉、藏语言的混合语——五屯话的混合性。

五屯话的动词体系表现为由基础来源语汉语的孤立类型向渗人来源语藏语的粘着类型变化的趋向。五屯话的动词有几十种形态变化。表示将做或未做行为的后缀[-dzi]、表示过去式后缀的[-lio]、表示完成式后缀的[-guo],这些时体后缀大都由汉语时体助词"着"、"了"、"过" 演变而来。表示领属关系的后缀[-tə]和表示处所关系的后缀[-li],也分别来自汉语的"的"和"里"。汉语的动词虽然也借助时体助词表达动作行为的状态,但是,无论从使用规模还是思维方式上,都与五屯话有着本质的差异[1]。五屯话的语法结构表现出与汉语完全不同的特征,最突出的是SOV语序,如"我红糖没有,我白糖有有"[ŋa xithie 'mie jɤ, ŋa phathie 'jɤjɤ]。与青海汉语方言里SOV虽较为普遍但也可以是SVO不同,五屯话一无例外地只有SOV结构。

与汉语语法手段相比,五屯话的显著特点除了具有类似

[1] 陈乃雄《五屯话的动词形态》,《民族语文》1989年第6期。

"格"的语法形式之外，还有与保安语也很相似的"数"的语法形式和祈使句的词尾。依据五屯话与原居民土族人所使用的保安语的相似之处，可以推测五屯话形成过程中接受了阿尔泰语的影响。当地的土族人（自称蒙古尔，又称察罕蒙古尔即白蒙古）与五屯人在历史上曾同属一个千夫管辖，来往关系十分密切，因此五屯话不可避免地受到原居民土族使用的保安语的渗透，"数"的语法形式和祈使句的词尾可能就是毗邻土族语渗透五屯话所留下的痕迹。贾晞儒说："（五屯人）现在讲的语言是夹杂着汉语词又类似阿尔泰语系保安话的一种特殊语言。"[1] 在对五屯话和保安语作全面比较研究以后，也许可以说：五屯话是汉语、藏语混合且受到阿尔泰语系保安语渗透的混合语。

（二）卡卓语：白语和彝语的混合

卡卓（"嘎卓"）语是现在云南通海县（由旧治通海、河西两县合并）蒙古族后裔所使用的语言，因其居民自称"卡卓"（$Kha^{55}$ $tso^{31}$）而得名。早在南诏、大理国时期，这一带已是白族和彝族的势力范围。1253 年，蒙古军队西征进入云南消灭大理后，河西地区成为蒙古军队的屯驻要地。在白族居民占多数、蒙古女性很少的环境中，蒙古士卒只有与当地白族妇女通婚，白语成为蒙古儿童的母语的同时，蒙古士卒在双语之后也就自然地逐步换用了白语，本族蒙古语趋于消亡。明初，政府派迁大批汉族人到河西地区屯田，当地的白族居民逐渐融合到汉族中去。只有居住在通海的部分操白语的蒙古族居民，没有与派迁而来的汉族同化。但他们却在与当地彝族居民的长期接触中，特别是与彝族的通婚中，所使用的白语受到彝语的强烈影响，从而形成了亦白亦彝的卡卓语。

---

1 贾晞儒《青海汉话与少数民族语言》，《民族语文》1991 年第 5 期。

根据和即仁的研究①,卡卓语与白语在语音方面保留着较多的一致性。卡卓语和白语都只有23个辅音,塞音和塞擦音不分清浊,没有舌尖后塞擦音;卡卓语和白语有着音值基本相同的16个元音。但是,卡卓语在白语4种音节结构类型之外,却增加了一种源于彝语的音节结构类型,即[m]、[n]可以自成音节,如"做"[$m^{33}$]、"牛"[$n^{33}$]等。此外,受彝语的影响,卡卓语音节以单元音为主,复元音主要出现在汉语借词里。通过821个常用词的比较,卡卓语与白语的同源比例占24.8%;通过850个常用词的比较,卡卓语与彝语的同源比例占44.2%。卡卓语的语序和彝语的语序基本相同,都是SOV式,而白语的语序是SVO式。以上这些比较证明,彝语对卡卓语的词汇和语序影响较大。

作为混合语的卡卓语,虽然在语音层面上保留了白语语音基层,但在词汇与语序层面上却被彝语所覆盖。由此可以推论,白语为卡卓语的基础来源语,而彝语为卡卓语的渗入来源语;换而言之,卡卓语的底层是白语,表层是彝语。然而,卡卓语既然是云南蒙古族使用的语言,其中是否有元代蒙古语残留的蛛丝马迹,还有待深入研究。据陈乃雄的研究,同样生活在云南的契丹后裔本人所说的"本话"还有个别契丹词语的遗存。②

(三)回辉话:占语和汉语的混合

宋元以来,由中亚或西域迁入汉地的波斯、阿拉伯、回纥的穆斯林,起初使用各自的本族语,后来逐渐换用汉语。但是,只有海南三亚市回辉、回新两乡的回族居民,其内部却使用一种特殊的回辉话。三亚回族是10世纪末以后陆续从东南亚的占城

① 和即仁《云南蒙古族语言及其系属问题》,《民族语文》1989年第5期。

② 陈乃雄《云南的契丹族后裔和契丹字遗存》,《民族语文》1994年第6期。

迁来海南岛的，最晚一批是在明成化二十二年(1486)，至今也已有500年的历史了。1941年，白保罗(P. K. Benedict)提出，海南岛的回民语言与占语有关[1]。欧阳觉非、郑贻青仅把回辉话同一些壮侗语从类型上进行比较，在没有把回辉话与其曾居地的占语比较的情况下，就否定了白保罗的观点，而认定与壮侗语近似[2]。文章发表后，回辉话的研究再度引起海外学者的关注。奥德里古尔(A. G. Haudricourt)与白保罗相继认为，回辉话与占语具有明显的对应关系[3]。郑贻青将回辉话与占语群的Rhade语作了比较，表明回辉话与拉德语有亲缘关系，但不能忽视回辉话在壮侗语的影响下，类型上发生了变化[4]。进一步的深入研究表明，占语只是回辉话的基础来源语，回辉话是占语和汉语的混合语。

占语和印尼语同源，属于南岛语系，其单词的音节结构是双音节结构。虽然现在的回辉话单词已经是单音节结构了，但以占语作为基础来源语的回辉话的前身无疑是双音节语言。在回辉人占语的单音节化过程中，回辉话中产生了补偿性的声调。回辉话和印尼语之间的同源词，对应部分多数在印尼语的第二个音节。如“田”[ma](回辉)—[hu-ma](印尼)。研究者由此认为，源于占语的双音节词，在周围单音节语(汉语方言、黎语)

---

① [美]本尼迪克特 *A Cham Colony on the Island of Hainan*，哈佛《亚洲研究学报》1941年。

② 欧阳觉非、郑贻青《海南岛崖县回族的回辉话》，《民族语文》1983年第1期。

③ [法]奥德里古尔《海南岛几种语言的声调》，《民族语文》1984年第4期。[美]本尼迪克特《澳泰语的相似语言：海南一种有声调的占系语言》，汉文译本见《民族语言研究情报资料集》第9期。

④ 郑贻青《再谈回辉话的地位问题》，《民族语文》1986年6期。

的影响下，双音节词的前一个音节脱落后便转变为单音节。这使人联想到沙加尔（L. Sagart）的研究[①]，原始南岛语发展为上古汉语曾经经历从单词双音节结构到单音节结构的演变过程，如“白”* burak < * brak < * bak。促使这一演变的动因缘于语言混合，由占语到回辉话的单音节化正是基础源语占语和渗入来源语汉语相互混合的结果。

在语音方面，回辉话有一些与汉藏语言明显不同而与南岛语言相同的特点。如没有以元音开头的音节；元音分长短两类；韵尾[-i]、[-u]、[-n]后面带喉塞音[ʔ]；韵尾前面的主要元音是紧元音；辅音[ʔb]、[ʔd]，前面有先喉塞音等等。在词汇方面，回辉话是以单音节词占优势的典型孤立语，只有少量的带[a-]、[i-]等前缀的双音或三音节词。回辉话词语的合成结构通常为NA顺行结构，如“旋风”[nin$^{33}$（风）vi$^{35}$（转）]，仅有少数合成词是逆行结构。NA顺行结构是南岛语的结构特点，少数AN合成词可能另有来源（比如汉语借词）。在语法方面，回辉话表现出较多的汉语特征，最明显的是利用虚词和语序来表示语法关系。虽然句子主干的语序和汉语基本相同，但回辉话保留着南岛语言定语放在中心词之后的特点，这与构词的NA相一致。

回辉话的研究存在一些“谜”。三亚回民来自占城，占城回民来自何方？北方回民大多是经由西域陆路而来，南方回民大多经由南洋海路而来，如果占城回民从阿拉伯经由海路而来，回辉话中是否还能发现阿拉伯语孑遗呢？现在的三亚回族居民会讲汉语海南话（闽南方言），会说军话（西南官话），还会说迈话与儋州话（粤方言的两种土语），回辉话究竟是占语和哪一种汉语

① ［法］沙加尔《汉语南岛语同源论》，第23届国际汉藏语言学会议论文，1990年。

方言的混合呢？另外，还有不少回民会说当地的黎语，在回辉话形成的过程中，黎语是否对之有所渗透呢？

郑贻青的研究进一步表明，回辉话的汉语借词及汉字读音与海南岛上的几种汉语方言读音有一定差别，通过回辉话借词及汉字读音与汉语海南话、儋州话、军话、迈话的对比，推断回辉话汉语借词及汉字读音是来自较古的汉语北方方言。郑文中所列回辉话51个汉语借词，其中“除夕”[sa:n$^{33}$ sa$^{55}$ za$^{11}$]当为汉语“三十夜”；“春节”[nin$^{11}$]当为汉语“年”。郑贻青认为：“回族子弟入学一直使用汉语汉文，其读音是一种接近军话的读书音，或称‘官音’、‘正音’。……当时海南岛北部的儋县（古儋州）的汉人除了使用着一种汉语南方方言之外，由于各类官员以及文人学士的传播，当地还同时通行一种北方方言的‘官话’读书音，……居住在这里的回族所学的很可能就是这种古‘官话’。”① 郑的研究信而有徵，有理由认为在回辉话形成的过程中，回族人带来的占语不是受海南岛南方汉语方言的影响，而是主要受历史上通行官话的影响。只是对于“当地还同时通行一种北方方言的官话读书音”，不宜含混地称为一种汉语北方方言。依据郑文所做比较，已经可以推定这种近古官话就是溯源于汴洛宋音、以南京音为标准的元明江淮官话。

## 三、历史混合语研究述评

随着对混合语存在事实的认可与语言混合研究的日益深入，语言污染论成为历史的偏见。在中国，虽然没有那种将少数民族居民学不好汉语斥之为糟蹋汉语的观念，但并不是说所有的语言学家都已经从源于欧洲殖民主义和沙文主义的语言糟蹋

① 郑贻青《回辉话中的汉语借词及汉字读音》，《民族语文》1995年第5期。

论或语言污染论的阴影中走出来。除了近代史上以西方殖民为背景引起的语言混合之外,有一个更为自然而古老的事实,在多民族国家或地区,语言的混合是特定历史条件下的语言互动,混合语是特定的民族杂居状况和特定的社会经济文化生活的历史性产物,语言混合是新语种产生的一种发生学模式。

依据从空间发现时间的原则,历史混合语的研究也是中国混合语研究的一个重要方面。1989 年,根据新石器时期三大考古文化系统与三大历史传说系统的相互印证,李葆嘉提出原始南方语、藏缅语、阿尔泰语融合而成原始华夏语的假说,从文化史角度预言了原始南方语和原始汉语的同源性①。沙加尔在第 23 届国际汉藏语言学会上发表的《汉语南岛语同源论》,运用历史比较语言学方法论证了汉语与南岛语的亲缘关系,可与之相互印证。1994 年,李葆嘉把语言的发生归纳为历时分化发生、共时聚合发生与克里奥耳语的混合发生三种类型②。刘丹青评价:"关于现有语种的来源,各教材基本沿用单讲分化的谱系学说,然而这早已难以面对语种多元发生的大量事实。《概论》将语种的发生分为……三种情况,具有更加全面的解释力"③。讨论汉语混合性质的还有吴安其的《十二兽古藏缅数词考》(1994)、陈其光的《汉语语源设想》(1996)等。混合语的存在事实为语言混合发生理论的建立提供了有力证据,语言混合发生的理论又为中国语言史和东亚语言史的研究注入了新的活力。

语言混合的研究有着十分深广的理论价值。传统历史比较

① 李葆嘉《试论原始华夏语的历史背景》(提纲),《语言学通讯》1990 年 1—2 合期。

② 李葆嘉《试论语言的发生学研究》,《南京师大学报》1994 年第 1 期;收入黄国营、余志鸿主编《语言学概论》,山西高校联合出版社 1994 年。

③ 刘丹青《改革、创新、飞跃与缺憾》,《汉语学习》1995 年第 6 期。

语言学无法解决的某些语言的系属问题，若能从语言混合的角度加以研究或许可望解决。所谓系属不明的朝鲜语，是属于阿尔泰语系还是别的语系莫衷一是。吴安其认为，朝鲜语有一批身体部位、天地日月等名称词、动作词和数词与南岛语词有对应关系，双方的语法框架又较接近，因此朝鲜语与南岛语关系密切。朝鲜语有关水稻文化的词语与南岛语词之间的对应，揭示了朝鲜先民原来是南岛语系的稻作民族，而不是阿尔泰语系的游牧民族[1]。是否可以这样设想，古代朝鲜语的基础来源语是南岛语，其混入来源语是阿尔泰语。也就是说，辽东—朝鲜半岛的早期居民是操原始夷越语（即原始南岛语）的古夷越人，与现代南岛人的祖先同一系统。远古时代，从朝鲜半岛、辽东半岛、山东半岛、江淮平原到江南湖泽丘陵一带，都是古夷越人的活动范围，以稻作为其主要生产方式。朝鲜族居住区域的北方是阿尔泰民族的世界，阿尔泰语对朝鲜语的影响至少可追溯至中国的秦汉时代。中古以后，契丹及契丹先民的语言对朝鲜语的影响长达 10 个世纪以上。阿尔泰居民进入辽东—朝鲜半岛，形成了民族的融合和语言的混合。古代夷越语（原始南岛语）和古代胡狄语（原始阿尔泰语）都是粘着类型，它们的后裔即现代南岛语和现代阿尔泰语仍然是粘着类型。因此，发生在辽东—朝鲜半岛历史上的语言混合是属于同一种语言类型而分属不同语言系属的两种语言的长期混合。主张朝鲜语归属阿尔泰语系的学者，仅仅注意到朝鲜语与阿尔泰语的类型相似性，而不知朝鲜语与南岛语的结构类型一致性和基本词汇的对应性。

依据沙加尔的汉语南岛语同源论，潘悟云认为回辉话可视

[1] 吴安其《论朝鲜语中的南岛语基本成分》，《民族语文》1994 年第 1 期。

为原始华—澳共同语向古代汉语演变的中间环节的参照之一[1],为汉语演变研究提供了新思路。李葆嘉从语言混合的角度探讨原始华夏语的起源,贯通汉南同源论、汉藏同源论与汉阿同源论,论证原始华夏语就是太古夷越语、氐羌语、胡狄语在中原地区长期互相渗透混成的结果。三种同源论在语言混合理论的阐释中得到化解,为揭开东亚南洋语言远古的面纱做了初步尝试,可以明确孤立语是语言混合的结果[2]。

## 四、混合语研究的价值与方法

与海外混合语产生的近代世界史上的殖民化背景不同,中国境内语言混合的社会背景是多民族的自然杂居、通婚生育。与从明代以来,中国南方存在过的广州洋泾浜英语和上海洋泾浜英语,基于当时中外贸易服务的市场需要也不同,中国境内民族语言的混合根植于长期的、共同的经济文化的日常接触交流。因此,西方混合语言研究中的以宗主国语言为"高层语",与以殖民地语言为"低层语"的概念以及混合语形成的单一起源论、克里奥耳连续流、语言简化理论、普遍语法理论、低层语影响论和综合理论等等,并不能机械地拿来阐释中国境内的混合语言和语言混合的现象。有些理论,比如单一起源论,不管是用以解释中国境内的语言混合,还是解释世界范围内的混合语言,明显是错误的。

就混合语的研究而言,有一些基本问题尚须探讨,语言渗透

① 潘悟云《对华澳语系假说的若干支持材料》,王士元主编《中国语言学报专题论丛系列之八·汉语的祖先》,1995年。

② 李葆嘉《华夏汉语三元混成发生论》,《建设中国文化语言学》,1994年12月;《论华夏汉语的混成发生和推移演变》,韩国《中国人文科学》1994年第13辑。

与语言混合如何界定与区分?混合语的认定标准是什么?中国境内语言混合中的共性规律和个性特征各是什么?语言渗透是语言混合的起点,但并不是所有的语言渗透必然导致语言混合。比如历史上的朝鲜语、日本语、越南语借入大量的汉语词汇,历史上的英语借入大量的法语词汇,以致这些语言都发生了深刻的变化,但是因为作为语主民族的这一群体的历史并没有中断,这些语言面貌的巨大变化通常被看作该语言的演变发展,而不是混合成为另外一种新的语种。

确定为混合语的语言,以现有的关系语言为参照,是一种既不属于A语言也不属于B语言的语种。一种情况下,操混合语的语主本身就具有混合性(卡卓人),该语主是通过民族融合形成的一个新民族,其语言自然是新语种。另一种情况,即使语主比较单一(五屯人),但由于长久脱离源民族的共同居住地域、共同经济生活与共同文化氛围,与源民族的经济方式、文化心理已经迥然不同,乃至源民族与原居地已趋模糊,需要专家考证。既然所使用的语言与源民族的语言已有本质差别,因此必须认为这是一种不同于基础来源语的独立语言。显而易见,一种受到多种语言渗透的语言是否确定为混合语,主要依据是语言的接触深度与语主的心理认同,其中后者至关紧要,因为语种的概念本质上是一个民族文化历史地理概念,而不是一个语言相异程度大小的概念。

语言学界有一种通行的说法,语音和词汇是语言的表层或浅层,语法是语言的底层或深层,并且认为表层容易发生变异,而深层语法结构则根深蒂固。喜饶嘉错批评了这种机械分层理论,"按一般层次观念来说,语音、词汇、语义、语法都各有其浅层次和深层次的内容,比如具体语音是浅层次的,但语音的规则、

系统、结构、框架则是深层次的，词汇、语义、语法也一样。"[1] 需要考虑的，这种语言各子系统皆分深浅层次的划分，是否也以稳定而不易发生变异与否为标准呢？根据系统论的观点，语言系统中的语音、词汇、语法之间是密切关联的。因此，任何一个子系统的变化都可能在不同程度上触发另外两个子系统的相关变化。各子系统内部也是密切关联的，如具体的语音和语音的结构、规则和系统是互为一体的。一种语言的语音系统应当包括音位系统与词音系统两个层面。抽象的音位系统寓于具体的词音系统之中。持续的词音系统层面上的具体语音要素的变化必然导致抽象的音位系统层面上的规则、结构以致整个系统的变化。通过对混合语的研究可以发现，语言混合过程中，语音系统、语法结构以及语言类型的变化，都是缘于语词的大量借入。也许，个别的借词不会产生影响或者其影响微乎其微，但是大量的借词不仅会使一种语言的词汇系统和语义系统发生改变，不仅会使一种语言的单词语音结构类型和构词方法发生改变，而且可以使借入语言的语音要素发生变化，由此引起音位组合规则、音节结构类型的改变，最终导致语音系统的本质变化。在一定的条件下，大量借词所带来的被借语言的构词方法以及语法范畴，可以引起借入语言的词法和句法结构的改变，最终导致语法系统以致语言类型的改变。因此，可以认为借词既是语言渗透的起点，又是语言混合的关键。中国民族混合语的研究表明，混合语中的语法结构往往与基础来源语差异较大，则更多地表现出混入来源语的特征，如上述五屯话、卡卓语和回辉话。基础来源语的语音系统一般较为稳定、不易变化，基础来源语的词汇所占比例可大（五屯话）可小（卡卓语），但混入来源语的语法往

① 喜饶嘉措《语言关系研究中的一些理论问题》，《民族语文》1991年第4期。

往覆盖基础来源语的语法。中国民族混合语的研究表明,语言的混合可以引起语言类型的转变。语言的混合可以从屈折语型变化为粘着语型(艾努语);语言的混合也可以由孤立语型转变为粘着语型(五屯话);语言的混合可以从粘着语型蜕变为孤立语型(回辉话)。后一情况表明,虽然孤立语型是语言混合的结果,但并非语言类型演变的终结。至于不同混合语的个性特征,就词汇而言,具有一些通用词、合璧词是五屯话的特点;由基础来源语的单词双音节结构变化为混入来源语的单词单音节结构是回辉话的特点;混入来源语的词语大面积覆盖基础来源语的词语是卡卓语的特点。

混合语言的研究在中国还是起步,混合语研究的理论和方法还处在趋于成熟之中。综合当前混合语研究的多个个案,这里尝试提出混合语言研究的五个步骤的系统方法。1.系统描写法。对假定性混合语的系统进行描写,归纳出该语言的主要特点。2.历史追溯法。追溯该语主的群体历史、追溯该语主的历史迁移线索、寻找该语主在历史上可能使用过的语言及其周边语言生态。以此为基础,初步确定其可能来源语。来源语或源语可以分为基础来源语与混入来源语。语主原先使用的语言为基础来源语,后来渗透其中或覆盖其上的语言为混入来源语。某一混合语的来源语可以是两种,也可以是多种。不同源语可以共时方式混合,也可以历时方式混合。3.源语比较法。将该语言与可能性来源语进行定性和定量比较。定性比较的结果,可能性来源语或者被确定为真实源语,或者被确定为虚假源语。如果出现后者,就需要重新寻找真实源语,再行比较。在一种语言有多种方言的情况下,应力求利用基础来源语的历史方言或参证后裔方言来与之比较,应避免利用非基础来源语方言进行比较。五屯人既然是明代由南京一带迁出,基础来源语当是明

代江淮官话，因此选用现代汉语普通话与五屯话比较，无论是从方言属性还是时间差异都不恰当，应当利用明代江淮官话材料、参证现代江淮方言与之比较。当然，在基础来源语方言系属不明的情况下，利用基础来源语的共同语进行比较也是难免的。定量比较则是在可能的情况下，分析计算出不同来源语在该混合语的语音、词汇和语法等方面的各自比例，比如卡卓语和白语、彝语之间的词汇比较。4.动态阐释法。对该混合语语言要素的历时混合过程，对该混合语的基本特征和结构类型，比如语言类型的转变进行阐释。5、理论抽象法。依据该混合语的研究引出相应的语言混合的理论。多种混合语的研究将有可能在几种次类型的基础上抽象出一个语言演化的模式，对语言的演变过程及其性质产生认识的飞跃，为历史语言学研究提供新的理论。

混合语言的研究和语言混合的研究前景非常广阔，深入研究中国境内及其毗邻区域中的混合的语言和语言的混合，创建中国混合语言学或东亚混合语言学以丰富世界克里奥耳学，是当代中国语言学界的一项历史使命。

# 第二章　人类文字的起源和演变规律

## 第一节　人类文字起源的多元发生论

根据考古发现，人类最古老的文字出现在公元前第四千纪左右，距今约有6000年。如果将成熟文字之前的一些单个符号视为文字的萌芽，由此追溯至公元前第六千纪年，文字史充其量8000年左右。语音一发即逝，祖先的声音不复存在。然而文字却以书面形式留传下来，为文字起源的研究，自源文字的出现和成熟提供了对象。

### 一、文字起源的神创说与学术性假说

（一）文字起源的神创说

文明古国都有文字之神。楔形文字传说是命运之神那勃（Nebo）所造，埃及圣书字传说是知识之神托特（Thoth）所造，中国汉字传说是黄帝史官仓颉所造，中美玛雅字传说是日眼大神（Kinch Ahau）的发明。古印度将婆罗米文字的发明归功于天神（Brahm），而将佉卢文字的发明归功于驴唇仙人（Kharofstha）。日耳曼人把鲁纳（Runa）字母的发明归功于斯堪的那维亚半岛神话中的最高之神奥丁。爱尔兰的欧甘字，“欧甘”（Ogham）即

是一位神话英雄的名字。神创说反映了古人将文字视为神圣之物,"鲁纳"一词即为"神秘"之意,视为人类智慧难以企及的创造。古希腊人虽然不了解人类文字的真正起源,但是他们知道字母的来源。这在古希腊历史学家希罗多德的著作中可以看到:腓尼基王子卡德摩斯来到希腊,带来了早先希腊人不知道的字母,于是卡德摩斯便被奉为希腊文字的创造神。

(二)古文字的发现和释读

1.尼罗河埃及文字的发现。1799年,拿破仑部下的一个军官在埃及发现了并列刻着埃及碑铭体、世俗体和希腊文三种文字的罗塞塔石碑。1822年,法国学者商博良(J. F. Champllion)释读了这种在古希腊称为"神圣雕刻"或"圣书"的象形文字。

2.西亚楔形文字的发现。1765年,德国学者尼布尔(C. Niebuhr)在波斯的波利斯城址发现了刻有楔形文字的铭文,明特尔(F. Münter)证实铭文中有一种为古波斯文。1802年,德国人格罗特芬德(G. F. Grotefend)释读出其中几个国王的名字。1835年,英国人罗林森(H. C. Rawlinson)在波斯发现了用3种楔形文字刻写的贝希斯顿摩崖铭文,并于1847年释读出古波斯文。经过许多学者努力又释读出巴比伦文,在20世纪上半叶释读成功埃兰文。1877年,法国考古学家萨尔泽克(E. de Sarzec)等人发掘伊拉克境内的拉格什城址,发现了前2500年的雕像和泥板,使人们第一次接触到苏美尔文字。在古波斯楔形字释读成功的基础上,释读了闪米特语的阿卡德文字,最后利用阿卡德文字释读了与闪含语系和印欧语系都无关的苏美尔语言文字。

3.中美玛雅文字的发现。16世纪,西班牙人进入中美,玛雅文字手抄本被当作多神教异物加以焚烧而由此湮没。19世纪,金斯葆罗夫(E. K. Kingsborough)在《墨西哥的古代文物》

中发表了仅存的三种玛雅文抄本,西班牙传教士迪耶戈·德·朗达(Diego de Landa)写于1566年的《尤卡坦记事》也得以出版。20世纪50年代,前苏联学者Ю·В·克诺罗佐夫将玛雅文字的性质确定为音节文字。

4.中国甲骨文的发现。1898年,天津孟定生、王襄见到了甲骨文。1899年,王懿荣鉴定为殷商甲骨文,人们得以窥见中国最早的文字。目前已释读了近2000字。

5.克里特岛文字的发现。1900年,英国学者伊文斯(A. Evans)发掘克里特岛克诺索斯王宫遗址,发现了象形文字和线形文字。意大利考古学家在挖掘费斯托斯王宫时,发现了费斯托斯圆形泥板,上面有另一种克里特象形文字。1952年,英国学者文特里斯(M. Ventris)和查德威克(J. Chadwick)成功地释读了记录古代希腊语的线形文字B。

6.印度河哈拉巴文字的发现。20世纪20年代,英国考古学家马歇尔(J. H. Marshall)主持印度河流域的哈拉巴文化的挖掘,发现了前2350—1750年之间的约有500多个符号的印章文字。

早期人类文字的发现为文字的发生研究提供了大量材料,揭开了远古历史的神秘面纱。

(三)文字起源的三种假说

1.图画记事起源说。由于古老的文字中都有象形字符乃至图画性极强,再联系到一些没有文字的民族往往运用图画记事,因此最自然不过的直觉是文字起源于图画。埃及圣书字的图画性非常明显。在古埃及语中,"画"和"写"、"绘画"与"文字"、"艺术家"与"书写人"这几组词分别相同。埃及学家M·A·科罗斯托夫在《埃及的语文学》中认为:"埃及文字的图画的、造型的性质,以及对其他民族象形文字体系的研究,表明了图画文字是象

形文字的出发点。"[1] 中国也有书画同源的传统说法。甲骨和金文中的字符象形较多，传统六书以象形为首。许慎《说文解字叙》："仓颉之初作书，盖依类象形，故谓之文"，"文"之本义即交错画也，或以为文即纹身之"纹"，以颜色涂抹于身体之上。许慎又云："文者，物象之本"、"书者，如也"和"象形者，画成其物，随体诘诎"，皆图画象形之说。

2. 刻划记号起源说。W·D·纳克在《文字起源》中说："考古学家曾在崖石上、石块上发现有史以前的人所描写的或镂刻的符号或字体，这些符号都很难解释。它们不是图画，而是不知属于何种系统的字形，……这些符号大概是原始野蛮人用来表示他们所有权的记号。"[2] 伊斯特林认为，古代的约定符号（还有图腾、巫术、工艺、计数等符号），除了个别的记号被吸收为词符字，从来没有成为词符文字赖以产生的基础[3]。

记号起源说在中国的影响即汉字起源于契刻说。于省吾认为契刻"还要早于其他的象形文字，这对于考证文字发生的萌芽状态，是具有重要关系的"[4]。郭沫若认为"中国文字的起源应当归纳为指事与象形两个系统，指事系统应当发生于象形系统之前"[5]。所谓指事，即契刻记号，郭文中又称为"随意刻划"。从而，一些人把半坡彩陶上的刻划符号肯定为中国文字的起源，或者中国原始文字的孑遗。裘锡圭认为，半坡类型的几何形符号不能跟古汉字里偶然同形的词符相比附，这两种符号显然是

① ［苏俄］伊斯特林（В. А. Истрин）《文字的产生和发展》（左少兴译），110 页，北京大学出版社 1987 年。

② 转引自王凤阳《汉字学》，42 页，吉林文史出版社 1989 年。

③ 伊斯特林《文字的产生和发展》，93 页。

④ 于省吾《商周金文录遗·序》，1957 年。

⑤ 郭沫若《古代文字之辩证的发展》，《考古》1972 年第 3 期。

不同系统的东西。文字形成过程中，通常会有一些记号被吸收为词符，但文字不起源于记号[1]。王凤阳认为，随意刻划没有社会性，带有很大个人主观性，刻划符号与文字的关系只是吸收和被吸收的关系，而不是源流关系[2]。

3. 原始记事方法起源说。汪宁生认为，文字是物件记事、符号记事和图画记事三类方法引导出来的，而不是仅仅起源于图画[3]。文字的起源至少与两个因素有关：一是关于文字符号的来由，一是关于文字原理的来由，以往的研究没有明确这两种区别。主张文字起源于图画或刻划记号的，其注意力集中在符号由来，而主张文字起源于原始记事的，则更多地注意文字原理的成熟。当然符号样式也包含一定的记事方法，但是主张汉字起源于契刻的，除了持有契刻比图画更具抽象性，更多的是注重于刻划符号与个别字符的相似性，却忽视了刻划符号是一个没有能产性的封闭系统，最大用途就是计数。图画符号具有能产性而且广义图画包括刻划，这好比写真画与抽象画。许慎"独体为文"，即包括象形、指事。于省吾在论及契刻早于象形文字时，也认为"以□为方，以○为圆"是"极符合于客观事物的真象"[4]。文字起源于刻划说的最大失误之处是把文字符号的起源与文字系统的起源割裂开来，把一些不成文句的刻划符号认定为"中国原始文字的孑遗"。文字符号之所以称为文字，就是因为它处于系统之中，哪怕是不成熟的系统。反之，处于系统之外的"随意刻划"何孑可遗！原始记事方法的考察，为文字的起源研究提供

[1] 裘锡圭《文字学概要》，23页，商务印书馆1990年。

[2] 王凤阳《文字学》，298页。

[3] 汪宁生《从原始记事到文字的发明》，《考古学报》1981年第1期。

[4] 于省吾《商周金文录遗·序》。

了参考，但是决不要以为原始记事方式必然导致文字的发生。

文字的起源研究或文字的发生学研究，需要探讨的主要内容应当包括：1. 没有发明或发明文字的民族在文字产生以前，所使用的原始记事方法的特征及其与文字发生的关系；2. 处于图画记事与表词文字之间的图形——语句文字有什么特点；3. 几种独立形成的自源文字——尤其是最早的苏美尔文字——是由什么演变而来；4. 人类文字产生的社会背景以及特殊触发因素。

## 二、原始记事方法与文字原理的成熟

口头语言以可听性载体来传情达意，与之不同，原始记事方法是以可视性载体来保存或传递信息。汪宁生将原始记事方法分为物件记事、符号记事和图画记事[①]。凡一事物被用来寄寓特定含义即为符号，文字及用于原始记事的实物、契刻、图画皆为符号。因此，原始记事方法依据符号形体特征可分为三类：实物记事、契刻记事和图画记事。原始记事方法并不是语言的派生物，它们的功能是记录事件，而不是记录语言，虽然这一事件可以用语言来说明。

实物记事方法有个别具象与一般抽象的区别。北美印第安人送"短斧"以示宣战，对方从地上拿起即意味着应战；用篝火、烽烟、鼓声来传递敌方人侵的消息。这些以个别物件为载体，通过借喻或约定来传递信息是个别具象实物记事方法。所负载信息单一但具有可视性，而可视性正是文字的首要属性。一般实物记事方法的典型是结珠和结绳。易洛魁人把代表不同观念的各色贝壳小珠串在绳子上。或以为算盘可能源于结珠，中国古文献中记载"上古结绳而治"。甲骨文中有些特别的符号类似印

① 汪宁生《从原始记事到文字的发明》。

加的结绳架,可能与结绳有关[1]。金文里的有些数目字则可能是古代结绳的孑遗[2]。据史载和民俗调查,中国的鲜卑、藏族、高山族、哈尼族以及环太平洋的琉球、夏威夷、美洲等地区都广泛使用过结绳记事,最为充分发展的是印加"结绳文字"。西班牙历史学家G·德·拉·维嘉(其母出身于印加部落),在17世纪对印加结绳作了描写:"为了表示战争、管理、贡赋、仪式,有各种不同的结绳,每一种结绳有许多纽结和系着的细线——红线、绿线、蓝线、白线等。正像我们用不同方式把西班牙语的24个字母加以配比,来表示不同的语音并且能分清差异一样,印第安人利用绳结的不同排列和颜色就有了大量的表示不同意义的方式。"[3] 由此可见,结绳记事方法已相当高级。在完善的结绳符号系统中,记录对象已经抽象化(绳结),记录单位已经单元化(以不同的绳结、以不同的颜色表示不同的事件、不同的性质),记录规则已经多样化(通过不同绳结的组合表示复杂的事件)。由于结绳系统中已隐含着符号的聚合性和组合性并具有生成性,因而就具有了信息容量增大的特性和记录较为复杂信息的功能,而单位性、能产性和组合性正是文字的基本原理。

契刻也是一种较为广泛使用的原始记事方法。许多民族都曾用木棍刻痕的方式来计算、建立契约关系或传递信息。斯拉夫各部落的"标签契约"——在木棍上刻上代表借出实物数量的痕记而一劈为二,债权与债务各执其半,这种标签在英、法和中国也曾使用过。大洋洲和非洲的一些部落有过一种讯木,出使者携带刻有痕记的木片前往目的地。这一刻木除了作为"书

① 曾宪通《汉字起源的探索》,《中国语言学报》第4期。
② 徐中舒《结绳遗俗考》,《说文月刊》1944年第4卷。
③ 转引自伊斯特林《文字的产生和发展》,93页。

信"，也还兼做信使的身份证。中国古代的符节与之类似。纳西族在木建筑上通常有三类刻划记号：计数符号、方位符号及占有符号，这些刻划符号只具有记事功能[①]。

《尚书·序》："古者伏羲氏之王天下也，始画八卦、造书契，以代结绳之政。"中国古代的八卦与契刻有关，其基本符号是"—"（连）与"- -"（断），取象于事物的分合互补，并以之组合演示事物的变化。契刻是具体的实物记事，而八卦是抽象的事理概括，是契刻式的古代哲学。由于3条横线最多排出8个卦形，卦形相叠演绎为64个变形，表达方式的高度抽象性不能反映语言的丰富多样性，因此八卦不可能成为文字体系产生的主体。古人把文字又称之为"书契"，"书"是以软笔涂写（繁体作書，从聿，者声；初文聿，即以又持笔之形）；"契"是以硬器刻写（初文从刀、从丰，像以刀刻划竹木以记事形）。此是以写字之法命名文字，并非文字起源于契刻之证据。

一些民族在20世纪还使用刻木记事。1928年，中缅边境的景颇族代表到当时的界务委员会请愿，拿出刻有11刻道的木块，代表11个当地最大的山官，表示他们誓属中国之意。50年代初期，中央慰问团曾收到一件傈僳族传信木刻。大意是："你们派来的三个代表在月圆时和我们相会，送上三包土产，分呈大中小三位领导。"这一刻木表明，要记录复杂信息，契刻的符号不能仅仅是记号，必须扩展至图形（O表示月圆）。虽然欧甘文字以刻线为字母符号，但这是在了解到拉丁字母原理后，才给本土源于手指和脚趾的计数符号赋于音值而制订出来的文字，并不能证明契刻记事可以自发导致文字的发生。尽管一些数目或指事字符可能取之于较早的契刻记事符号，但契刻记事与文字体

① 宋兆麟《纳西族的刻划符号》，《化石》1981年第4期。

制没有传承关系。然而,契刻记事表明人们已经可以摆脱具体物件,掌握了运用平面线条记录信息的技能。与实物记事方式相比,契刻记事已由立体转为平面,由实物转为刻线,而平面性与线条性也是文字的基本属性。

实物记事与契刻记事的同时,图画记事的方法逐步从原始绘画中分离出来。属于欧洲旧石器时代晚期的奥瑞纳—索留特列亚文化(前 40,000 ~25,000)和马格德林文化(前 25,000 ~15,000)的岩画,都是不具有情节叙述的写实性图画,因此还不能认为是记事性图画。在中国境内也发现了许多岩画,如天山以北岩画、青海哈龙沟岩画、甘山黑山岩画、内蒙白岔河岩画、大兴安岭岩画、广西花山崖画、云南沧源岩画、四川珙县岩画、江苏连云港将军崖岩画。因为这些岩画中的图形没有语言单位化,组构方式与语言表述并不一致,所以图画文字的推测可以排除。但是,到底是艺术创作还是图画记事却颇有争议。如果认为图画记事必须要有事件主体,而且要以叙述方式展示出来,那么这些岩画应确定为艺术。从历史演变的观点来看,原始绘画艺术与图画记事之间有一个模糊阶段,换而言之,原始图画记事只是原始绘画艺术的一种特殊应用。极北地区和美洲土著留下了一些生动的图画记事,较有名的如《爱斯基摩人的游猎日记》、《大科塔人的编年史》、《印地安海狸换野牛的契约》和《奥杰布娃的情书》等。图画记事方法在这些没有文字的民族中应用极广。

图画记事已具有书写的性质和图形的特征,因此有人称之为"文字画"(近似文字的图画),有人称之为"图画文字"(图画式的文字)。前者强调它的图画性质,后者强调它的书写性质。虽然可以用或者必须用语句解读这些图画,但由于图形没有与固定的语言片断配比,即图形未语言单位化,在解读中具有很大的随意性。尽管有极个别图形已近乎表词,但总体上仍然是图画

记事的手段。因此，孕育着象形文字符号的图画记事尚不能称之为文字。近年来在山东大汶口发现了年代约在前2000～2500年之间的陶器符号，文字学家以为这为探索中国文字的形成投射了光明，然而在文化史家并不确认为文字。研究历法文化的庞朴，以为远古曾用大火为示时天象，此乃日月会于大火之象[①]。在研究生殖文化的赵国华，却看成象征男根的山纹与女阴纹的组合之象征，寄托了先民生殖繁盛的祈求[②]。图形符号的解释，随着研究者的视角而出现万花筒般的变化。

原始记事方法的研究显示，实物记事表明人们已具有用可视性、可保存性的载体储存、传递信息的能力；结绳方法中已包含着用有限基本符号加以排列组合储存复杂信息的原理，即符号的单位性、能产性和组合性基本原理；契刻记事使符号平面化、线条化；图画记事以物象代表物体，使记事方法具有书写性。以上这些：符号的可视性、储存性、单位性、能产性、组合性和平面化、线条化、书写化，都为文字创制原理的成熟提供了准备，而图形及一些刻划记号则可能成为最初的文字符形。然而，原始记事方法并非必然演变为文字系统。恰恰相反，正是由于没有发明或仿制文字，一些民族的原始记事方法才保留至今。

## 三、从句符文字的特点透视文字发生

既然文字定义为表现语言的符号系统，那么可以从符号的形状和功能两方面对文字分类。根据符号形状，文字可分为图

---

① 庞朴《火历钩沉——一个遗失已久的古历之发现》，《中国文化》1989年创刊号。

② 赵国华《生殖崇拜文化论》，292页，中国社会科学出版社1990年。

符、文符与标符;根据符号功能,文字可分为表语句、表语词与表语音。从符号的形状(能指)与功能(所指)双向界定,文字主要有三种类型:图符——语句文字即句符文字;文符——语词文字即词符文字;标符——语音文字即音符文字。前一种类型是不成熟的文字,后两种类型是成熟的文字。

原始记事方法与文字系统的区别就在于记事还是记言,当图画记事用来记言时,图形就变成了句符文字,通过对不成熟的句符文字的形成过程的研究,可以窥见图画记事如何突变为图符记言。单个图形与单个语词相对应时,可以判断这是文字,如《奥连的墓碑》中画着一头倒鹿。"鹿"在印地安语中叫"奥连",这位头领也叫"奥连",做人名符号的"鹿"的图形就成了文字符号。倒过来的鹿形不仅表示人名,而且表示"死去",表明图画记事中萌发出个别表词字。

因为图画记事也可以用语句解释,因此判断一连串图画是记事还是记言时会左右为难。从结构来看,记事性图形以整体呈现,图形内的各部分之间的摆布漫无规则,符号未语言单位化。而记言性图形以分节呈现,图形内的各部分之间以线性方式较有序排列,符号已语言单位化,表现的往往是语句。不杂用词符的句符文字是原始的,而杂用词符或以词符表句义的句符文字是发展之中的。问题在于,如何判断图画记事和原始的图形记言。印地安人中一些记录歌词的图形,以勿庸置疑的方式告诉我们:这是在记言。比如印地安人的"巫歌"[①],每一图形相当于一句歌词,此外还有"战歌"以及"情歌"[②],为我们找到了图

① 伊斯特林《文字的产生和发展》,60页。

② 葛劳德《比较文字学概论》(林祝译),商务印书馆1930年。转引自王凤阳《文字学》。

画记事突变为图形记言的某种明证。

中国有两种典型的象形文字:沙巴文和东巴文,它们都是用来记录宗教经文的较为成熟的句符文字。经文用于口诵,为记录语言的文字确定无疑。其中的图符一般都有较为固定的音和义,甚至出现假借。但是这些图符在文章中所表达的是围绕这个图符的语句含义。这些图符字具有两种性质:一是在行文中的语句含义,一是在行文外的语词含义。在备用状态下,图符字是表词字;在使用状态下,图符字是表句字,从而显露出由句符文字向词符文字演变的端倪。沙巴文中图符排列的非线性,东巴文中的引导字以及倾斜、倒置、填色等附加语义的方式,都是图画记事或初期图形记言的胎记。句符文字尽管是以图符模糊地对应语句或语词,而未能准确反映句子内部的语词和结构,但是它是以表现语言为功能的原始文字而不是记事图画。

## 四、苏美尔文字起源于粘土标志说

文字起源的假说,不仅借助于原始记事方法的参证,更重要的是依据对自源文字形成过程的探索。文字的发生学研究,其核心问题即自源文字的起源研究。通过对图符——语句文字的产生和发展的研究,可以推定文字的来源是图画记事,埃及圣书字、甲骨文与玛雅文的象形性都支持着这一假说。然而文字起源于图画的假说,在世界上最为古老的文字——产生于公元前40世纪的苏美尔文字的起源研究中遇到了挑战。美国德克萨斯大学的斯迈特·彼塞特(Denise Scbmanbt Besserat,1978)在《文字的始祖》中提出"苏美尔文字起源于粘土标志"的假说[①]。

考古发现,西亚地区有一个在公元前第十千纪就出现的用

① 载[美]王士元主编《语言与人类交际》,广西教育出版社1987年。

于清算账目的粘土标志系统,即实物记事的一种方法。这一系统大约有15种主要类别的粘土标志,这些标志又可根据大小、制作、色泽及组合的不同,分为大约200个小类,分别代表不同的数目与实物。而在两河流域古城乌鲁克(Uruk)所发现的苏美尔泥板上的许多表意文字,几乎是在平面上精确地复制了这些粘土标志。不仅表示数字的压印,就是表示羊、衣服、金属等的泥板字符,都有相应的粘土标志。这个古老的记录系统,出现在西亚最早的农业居民点建立的新石器时代早期,粮食的贮藏与交换有可能触发了这一系统的出现,并在西亚延续了五六千年。到了青铜时代早期(公元前第四千纪),农业有了长足的进步,手工业专门化且批量生产。以贸易为主的城市经济的发展对记录系统有了新的要求,不仅产品就是存货、载货,工资也必须记录。复杂的商业会计引起传统粘土标志系统的变革。起初出现了如同提货单的"泥泡",生产者在把货物委托给中间人时,交给他一个里面放着表示货物类别和数量的粘土标志而且外面加盖印章的泥泡,接收者打开泥泡就可据标志核实货物。逐步发现可以像加盖印章一样,在泥泡表面压印上泥泡内所装粘土标志的形状和数量,不必打开就能"读出"泥泡内装着什么。曾经发现过一个蛋形泥泡书版,表面的印文旋转着看是一张列有48种牲口的清单,而里面恰巧装有48个粘土标志,一份书面清单和一批可数的粘土标志结合在一起。再以后,就是用粘土标志的平面压印来代替立体标志本身——原始记录系统突变为文字系统的关键——泥版上的苏美尔文字诞生了。

除了不少表意文字与粘土标志的平面压印之间存在着明确的一致性以外,选择粘土那种不适合的材料用作书写工具,也同样可以证明苏美尔文是从泥泡继承下来的。因此斯迈特·彼塞特认为,两河流域文字的最早用例,可能不像很多人所设想的那

样源于图画或纯粹发明的结果，而是苏美尔人对西亚古老记录系统的变革性应用。按照这一假设，由于城市的兴起和贸易的发展，传统的农业社会的记录系统被推到新的商业交换的轨道上来，粘土标志的压印形象很快取代了标志本身，这一突变导致了两河流域出现了人类文明史上最早的文字。

在这一假说之前，对苏美尔文字的起源进行过探索。根据德国弗尔根斯坦(A. Falkenstein)和苏俄 П·М·季亚康诺夫的看法，公元前 3200～3100 年的古苏美尔铭文符号有三类：第一类是数量不多的组合图形，前 3100 年以后开始消失；第二类是单个实物图形，存在到前 2500 年；第三类是占多数的图示符号，后来把前两类排挤出去。然而，对早期铭文的性质却有不同看法。弗尔根斯坦认为每个图形表示一个单词，因而是“词文字”。季亚康诺夫认为古苏美尔文字具有 4 个特点：1.符号并不表示构成信息的所有词，只表示关键性的词；2.每个符号可以表示不同的近义词和一个词不同的语法形式；3.符号排列顺序是随意的，而与词序中不一致；4.缺少萌芽状态的表音要素。由此他认为“苏美尔文字同图画文字没有原则区别”。伊斯特林认为，既然季亚康诺夫又承认每个符号“相当于一定的概念和词”，古苏美尔文字就应该看作表词文字。至于古苏美尔文字表达的只是关键词及排列无序性，并不动摇它的表词文字的基础，仅仅证明这个基础发展得还不够充分而已。季亚康诺夫所列举的古苏美尔文字的特点，与上文所分析的处于句符文字向词符文字过渡中的文字状态较为接近。至于上文把东巴文归入句符文字而不是词符文字，是因为它们图画性极强，真正的“词符”还很少，而沙巴文比东巴文更原始些。既然古苏美尔文“词符”已占多数，因此可定为“准词符文字”，从而可以推断苏美尔文字可能有过句符文字阶段。

关于句符文字或前表词文字阶段，有人认为在苏美尔人移居两河流域之前存在。即使接受了这一假说，仍然要解释是什么原因促使苏美尔人较快地过渡到表词文字。伊斯特林认为，这在很大程度上决定于苏美尔铭文的特殊内容。与古埃及历史宗教铭文不同，苏美尔文通常是生产计算的经济文件，这种格式特别便于把图画文字分解为单个的表词字——指明数量的数目字和指明计算实物的图画表词字①。但是，伊斯特林未能揭示经济文件中表词原则的来源，而斯迈特·彼塞特的“粘土标志说”解开了这个历史之谜。因此，合理的解释是：苏美尔人在移居西亚之前可能有着原始图画记事或不成熟的图符——语句文字，在对西亚古老记录系统的变革性应用中将二者结合。换而言之，表词原理与一部分表词字来自西亚粘土标志记录系统，另一部分图画性较强的符号来自原先的图画记事，二者的结合促使苏美尔文字较快地过渡到表词文字阶段。

尽管苏美尔文字可能起源于粘土标志记录系统，但是不能由此否定其他几种自源文字起源于图画记事方法。文字的起源并不限定于同一种原始记事方法，而是与文字发明者未有文字之前所使用的传统记事方法密切相关。国外一些文字史家所提出的词符文字源于苏美尔文字的单源论假说，在不同自源文字起源于不同原始记事方法的多元论面前难以成立。

## 五、文字发生的社会背景和触发需求

文字的发生有待于社会文明的发展和需求的触发。苏美尔时代是两河流域国家产生和发展的时代，也是铜石并用和青铜时代。欧贝德时期（前 4300～3500）的灌溉农业发达，养牛、种

---

① 伊斯特林《文字的产生和发展》，130 页。

植麦子和亚麻，盛行着广泛的铜石和木材的贸易。信仰母神的神庙已具有后代神庙的特征，一些男性雕像的出现表明向父权制过渡。以神庙为中心开始出现初期的城镇，乌鲁克时期（前3500）的大规模神庙建筑物反映了当时社会的高度统一和统治力量，城邦国家出现。制陶和采石等手工业分离出来，金属制品盛行，出现了文字（已发现2000多个）。杰姆代特奈斯尔时期（前3100），文字使用普遍，泥版文书主要记载经济活动，使用者主要是商业人员。

前第五千纪年的古埃及进入农业定居，用牛挽耕。前第四千纪年进入铜石并用时代。前3500年形成上下埃及王国，前3000年美尼斯统一埃及。虽然古王国时期（前2686～2181）的手工业迅速发展，内外贸易随之增进，但这并不是古埃及文字发生的社会因素。古埃及在前王朝时期（前3400～3100）已出现文字。纳尔迈石板上的象形文字，用“鱼叉”图形表示国王的名字“纳尔迈”。更早的是斯科尔皮昂王权标头上的浮雕，国王头部右侧两个圣书符号：象征国王的“七瓣花结”和国王名字的“蝎子”。统一后的埃及，其历史开始有文字记载可考，从第五王朝（前2500～2350）史官记录各朝国王的名字和按年发生的大事。古埃及文字的主要使用者为祭司，文字通常刻在庙墙、陵墓、石棺、雕像等宗教纪念物上，因此被称为“圣书”。

前13世纪，甲骨文已是一种相当成熟的表词文字。商代的社会经济相当发达，农业是当时主要的生产部门，手工业专业分工很细，冶铸青铜和纺织提花技艺很高，已用货币交换。人们信仰自然神和祖先神。掌握文字的是贞人，甲骨文绝大多数是刻在甲骨上的卜辞，与原始信仰密切相关。中美洲玛雅人以农业和狩猎为生，崇拜自然神。公元初尤卡坦半岛形成最早的城市，并发明了文字。玛雅人每隔几年要刻一块记录历史的石碑，现

存最早文字是328年的沙克敦城石碑上的铭文。只有祭司才识字,他们用玛雅文记录神话传说、编写祈祷文和编年史等[1]。

通观自源文字出现的前后历史背景,人类文字发生的一般社会条件是:1.农业社会经济;2.手工业专门化;3.使用自然铜或冶炼铜;4.商品交换发生;5.出现城邦国家;6.信仰原始宗教。然而,具备这些条件并不意味着一定出现书写型文字。南美印加人,其祖先早在公元前就掌握了制陶、建筑和冶炼技术。15世纪中叶的印加王国已相当强大,用青铜制造工具和兵器,棉毛织物绚丽多彩,进行产品交换。盛行祖先和太阳崇拜,国王被认为是太阳的化身而"印加"意为"太阳之子"。除了在每个市镇设有结绳官,还为贵族子弟设立了学习语言、祭礼、历法的学校,传授的是"基普"即结绳文字。虽然有过一些文字图画或原始图画记事,但并未由此演变成文字。即使在独立发生文字的古老文明中,文字的早期掌握者和应用场合的不同,也表明了文字的发明可能出于满足各自不同的需求。苏美尔文字的出现是商业会计的要求;圣书字、甲骨文和玛雅文主要掌握在祭司、贞人和巫官手中,用于记载宗教文辞和王朝历史,似乎精神信仰的需求成为触发文字发生的直接因素,在大科塔人的"编年史"、印地安人的"巫歌"中可以看到这种端倪。文字的起源是人类文明发展到一定历史阶段,新的经济运作或精神信仰的社会需求促使以前所用原始记事方法实现的变革性突破——由记录事件突变为记录语言。正是这种原有不足方法的中断,才促使新的方法得以产生,并迈上由不成熟到成熟的完善之路。

---

① 资料来源于周一良、吴于廑主编的《世界通史》,人民出版社1973年;《中国大百科全书·考古学》,中国大百科全书出版社1986年。

## 六、文明的多样性与文字的发生多元论

(一)人类的原始记事方法具有多样性

没有发明文字的民族或发明了文字的民族，在无文字之前所用的原始记事方法并不只有一种，而是多种。这就是原始记事方法的多元性。在不同的民族中，这些不同的原始记事方法的发展是不平衡的。西亚土著的粘土标志记录系统非常实用。印加人的图画记事并不充分，结绳方法却非常完美。原始记事方法为文字的发生提供了原理的和符形的可能条件，但是原始记事方法并不必然引导出文字系统。有着完善"结绳文字"的印加人并没有发明文字，也许正是因为结绳文字的完善和发达反而使其不思变革。

(二)人类文字并非源于一种原始记事方法

在古老的自源文字中，有的可能较多地起源于实物记事，有的可能较多地起源于图画记事。就是在一种文字中，文字产生的要素——原理和符形——也可能并非是单一的。人类最古老的苏美尔文字，虽然原理和符号主要源于西亚的粘土标志记录系统，但也可能融进了苏美尔文原有的图画记事符号。而起源于图画记事的中国古文字，在形成过程中也可能吸收了契刻记事符号。

(三)触发文字发生的需求因素的多样性

虽然这些发明文字的古老民族在进行这一伟大创举时，都已迈入铜石并用时代或临近青铜时代，并且出现城市和国家，但是，并非进入青铜时代、出现城市和国家就一定会产生文字，尽管他们已有相当完善的原始记事方法。即使那些创造了文字的民族，触发文字发生的直接因素也不尽相同，有些可能是为了满足物质生产的需要，有些可能是为了满足精神生产的需要，包括

源于自然神崇拜的对神的信仰，或源于祖先神崇拜的对历史的追溯，从而从另一方面显示了文字起源的多元性。由于人类文明的发展具有偶然性，人类文字的起源或发生必然呈现多元状态，人类文明之树开放出积淀着古老民族史前文明的五彩缤纷的文字之花。然而，尽管各种自源文字由来的原始记事方式可能不同，触发文字发生的直接社会因素可能不同，但是有一点是共同的——所有自源文字的出现都是对传统记事方法的变革性突破，这是人类文明发展的必然趋势。

## 第二节　自源文字的发展和借源文字的突变

人类文字的发展有没有共同的基本规律？这一规律是否应表述为"由象形文字到拼音文字是文字发展的基本规律"[①]？有哪一种文字的发展经历了象形——音意——拼音文字三个阶段？依据文字的发生学研究，作为人类进入文明阶段标志的文字，可以分为原创性的自源文字和仿造性的借源文字。原创性的自源文字是形体或原理在没有任何借鉴的情况下自发形成的文字，仿造性的借源文字是形体或原理有所借鉴的情况下自觉制订的文字。人类文字发展的一些规律，可以从自源文字和借源文字两个方面来加以深入探讨。

### 一、自源文字并未发展成为拼音文字

（一）苏美尔文字并未发展为拼音文字

不迟于公元前5000～4000年之间，苏美尔人已在两河流域定居，他们的语言同闪含语言没有亲属关系。根据研究，苏美尔

① 李兆同等主编《语言学导论》，203页，新疆大学出版社1981年

语属于粘着结构类型和具有做格特色的语言。

苏美尔文字是人类最早的文字。公元前4000年左右，苏美尔文字已进入相当发达时期。对于这种压写在泥板上的文字，后世阿拉伯人称之为钉头文字，而英国人则称之为楔形文字。苏美尔语大多数是单音节词。音节的构成方式一般有四种：V、VC、CV、CVC。起初的实物符号就是为单音节词设计的，假借的方式促使文字中出现了音节符号。“洋葱”的读音 su-mu 也是动词“给”的一个形式，于是“洋葱”的符号也用来记录动词“给”。要表示双音节词，就把两个单音节词加起来，这两个单音节词在双音节词中仅仅是音节符号。起初，这些音符与拼音文字中的音节字母还很不相同，但和汉语联绵词里的表音汉字的性质差不多。苏美尔文字符号达900个以上，大致可分为意符、音符和部首符号（近似类符）三类。由部首和音节构成的复合符号就是形声字。乌鲁卡基那时期（前2370）的苏美尔字符减到600以下，其中约有100个是表示V或VC的音节符号。苏美尔语的文字中没有单独表示辅音的符号，这与辅音占重要地位的古埃及语不同。虽然苏美尔文有大量的音节符号，但只有在某个词没有特殊符号的情况下才使用。直到公元前2006年被阿卡德人征服之时，苏美尔文依然还是意音文字。

（二）古埃及文字并未发展为拼音文字

公元前3500年，尼罗河流域的古埃及人已有完整的文字，与现代埃及人语言属于闪米特语族不同，古埃及人的语言属于含米特语族。古埃及文字约有几千个符号，常用的七百多个。这些符号分为意符、音符和类符3种类型。仅有描绘事物特征的意符是不能完整地记录语言的，这在东巴文中表现得很清楚。许多概念不便于用象形方式制造符号，而这些概念的名称可能与已有象形符号的某一事物的名称相同，于是古埃及文字中的

最初意符便通过假借方式具有了第二种功能——充当音符。"燕子"一词的发音与"大的"一词的发音都是 w-r，"燕子"一词的符号便成了"大的"一词的符号。进一步只要碰上 w-r 这个音组，就可能用"燕子"一词的符号来表示①。有一定语音的意符——临时借用的个别音符——固定使用的扩大音符，这就是音符产生和发展的轨迹。由于古埃及语中的辅音具有负担词根意义的特别作用，人们便根据离素原则创造了辅音音符，比如用"手"的符号表示辅音 d。在古埃及文字趋向成熟的过程中，除了二音符、三音符 75 个以外，在古王国时期（约前 3000）逐步固定为 24 个单音音符。这 24 个音符，既存在多符一音现象，又不能拼写出全部词汇，更没有形成字母表，因此与拼音文字中的字母有所区别，不宜称之为"辅音字母"。

有了意符和音符的文字，便能完整地记录语言。一方面，同音替代的结果出现了同音异义符号，符号的清晰度受到影响，势必引进类符来分化同音字而构成了表意形声字；另一方面，给意符加注音符的结果构成了注音形声字。随着偶然使用个别音符或类符，到有意识地利用音符和类符拼合形声字，某种文字最终进入形声阶段。依靠音符和类符的组合，符号的能产性便呈现出无限状态。因而，一种具备了形声方式的自源文字，既然记录语言功能的造字方式已经满足，也就无须在表意、表音、形声之外，再发明第四种方式了。造字方式完全具备的这种文字，除了形声的日益扩大，以后的变化主要表现为字体的变化（图画体与线条体、异体与正体、繁化与简化、手写体与印刷体）和字数的增减。古埃及自源文字经历了几千年的发展，在它死亡之前还是

① ［苏俄］科车尔基娜（B. A. Коуергина）《文字史概要》（杜松寿译），中国科学院语言研究所编《语言学论文选译》第 5 辑，中华书局 1958 年。

意音文字。

(三)自源文字都没有发展为拼音文字

在地中海文化圈,公元前 20 世纪初,操涅西特语的印欧人来到位于小亚细亚的哈里斯河流域,同化了当地非印欧语的土著,而成为历史上的"赫梯"。赫梯人原有的象形文字至今尚未释读,虽然赫梯强盛时所用的文字为巴比伦楔形文字,但传统的象形文字一直用于铭刻和印章,直至公元前 8 世纪赫梯被亚述所灭。公元前 2000～1700 年,地中海地区的克里特岛存在一种象形文字,前 1600 年左右演化为线性文字,至今尚未释读成功,可能都是意音文字。在印度河流域,哈拉巴考古文化(前2300～1750)已经有了文字,主要保存在石陶或象牙等制成的印章上,在陶器和金属器物上也有所发现。已经发现的 500 多个符号中包括意符和音符,据此可知是意音文字。

在中美地区,生活在墨西哥南部和于加丹半岛的玛雅人,在公元后最初几个世纪内创造了书写系统。除了刻在石碑、祭坛、庙宇墙壁上的铭文,也有许多写本。这种具有魔法性质的文字的释读工作还在继续进行。它们似乎是意符同音符的结合,其中有用一个意符来表示另一同音词的方法。在墨西哥北部生活的阿兹台克人的象形文字中,也有由截取方式产生的音节符号。

纳西族的东巴文是 20 世纪中叶还在使用的象形文字。至少在 11 世纪中叶,东巴文已用来写经。虽然纳西人后来又有了音节类型的哥巴文,但民间一直有人用东巴文来记帐和写信。东巴文约有 1400 个符号,象形字可以通过修饰(点、线、色)或变形(倾斜、倒文、缩减)的方式表明附加语义。东巴文中存在假借字,借用与"小"同音的"剪刀"的字符做"小"的字符,此外也出现了一些形声字。由于象形字占主导,假借、形声还在初步阶段,因此东巴文还不能完全以字词单位对应性地记录语言。经书中

有时只写重要的引导字，一个字可以读成一句话，几个字可以读成一段话[①]。只有假借和形声进一步发展的文字，语段或句符文字的痕迹才会被完全淘汰。

## 二、自源文字在历史上的命运

(一)文字随着文明夭折和语言死亡而湮没

公元前13世纪，赫梯王国在海上民族的打击下分裂，残存的几个小国到前8世纪为亚述人所灭。约前14世纪，克里特岛米诺斯王朝为来自希腊半岛的亚该亚人推翻，再以后是希腊半岛的多利亚人入侵，克里特文明湮没。印度哈拉巴文明毁灭的原因，通行说法是由于来自西北的雅利安人的入侵。在文明的夭折过程中，随着民族语言的死亡文字也湮没了。

与这些文明和文字同归于尽有所不同，作为古代西亚国际文字的楔形文字，在苏美尔文明毁灭后又相继被辗转借用了2000年左右，直到公元前后才彻底湮没。说闪米特语的阿卡德人，在征服了苏美尔的同时因袭了楔形文字。不久，苏美尔人取得独立，但又被阿摩利人与埃兰人征服，逐渐与闪米特人同化。虽然苏美尔语作为口语已弃而不用，但作为宗教文献语言又保存了相当一段时间，尽管到公元前第一千纪前半期，苏美尔语在文献中也不见了，但是苏美尔楔形文字式样却依然被辗转沿用。前1894～689年，随着与阿卡德语相近的阿摩利人建立的古巴比伦王国，楔形文字进入巴比伦时代。与绝大多数非拼音式借源文字一样，巴比伦文字主要靠音借和训借构成，即在用符号表示音节时按照苏美尔语的音读，但在表示概念时却改用闪米特语训读。一些音符是依据巴比伦自制的单音节词转化而成的，

① 《中国民族古文字研究》，中国社会科学出版社1984年。

同时也采用复合方式构成形声字，因此虽然音节符号有几百个，但巴比伦文总体上是意音文字。亚述人和建立新巴比伦的迦太基人都沿用了巴比伦文。前1595年左右，赫梯人洗劫巴比伦，与先前的闪米特人不同，赫梯人的涅西特语是印欧语。为了适应赫梯语的屈折特点，更多地采用音符来表示词尾变化，但其文字不是拼音文字，而是一种既用意符也用音符、许多词是用形声方式表示的意音文字。前5世纪，古波斯既有用于官方和商业文件的阿拉米字母文字，却又从埃兰人和米底人那里承袭了楔形传统，制订了波斯楔形文字，41个符号中有36个音符、4个表意符号和1个隔词符。

（二）异族入侵引起文字的被迫换用

公元前525年与前343年，波斯两度征服埃及，摧残了古埃及文明。前3世纪，亚历山大灭波斯，希腊人在埃及建立了托勒密王朝（前305～30），定希腊文为官方文字，并用20个（一说25个）希腊字母拼写埃及语，另外采用7个平民体音符来表达埃及语中的特殊发音，古埃及文字逐渐被科普特文字（希腊人称埃及人为"科普特"）所代替，同时古埃及语的最后阶段则被称为科普特语。阿拉伯人征服埃及后，714年定阿拉伯语为国语。随着古埃及人被同化，科普特语文的使用范围越来越小。16世纪的上埃及一些地方仍在使用，据说迄今仍残存5种方言，另外科普特教派宗教仪式时仍然使用这种语言文字。随着哥伦布发现新大陆，欧洲殖民者的侵占，玛雅人和阿兹台克人的象形文字被视为异教巫术废除、文献被焚烧以后，西班牙人引进了罗马字母，形成了尤蒂—阿茨蒂克文字与基切文字等。

（三）文化延续使用至今的自源文字

自源文字中能够延续至今的，只有导源于商代甲骨文、迄今已达3500年历史的华夏文字。但华夏文字并非迄今为止使用

历史最长的文字，在公元前后废弃不用的楔形文字，苏美尔人使用了2000年，作为古代西亚国际通用的文字，其他民族又辗转借用了2000年，前后存在了4000年。苏美尔和阿卡德、巴比伦等不是一个民族，商、周同样不是一个民族，由夏、夷等混合而成的汉族的心理认同感在汉代才最终形成。另外，虽然在20世纪中叶东巴文还有人使用，但纳西人早已有一套音节文字。东巴文不能完整地记录语言，它的自然消亡是在劫难逃。

## 三、在借用突变中产生拼音文字

尽管苏美尔文使用音节符号，埃及人发明了辅音音符，形声方式在汉字中充分发展，但是表音方法只能作为一种手段在意音文字中存在，它无意冲破意音体系的藩篱。从意符到音符，再到形声，实现了哲学中的"正——反——合"，宣告了自源文字没有可能自发演变为拼音文字。只有当一个没有文字的蛮族，往往是作为征服者来到一个文明程度较高、有着文字的民族面前时——文明程度较低的民族用武力征服了文明程度较高的民族，文明程度较高的民族文化同化了文明程度较低的民族文化，征服者被征服了——或许，他们会彻底解放表音方法，通过借源方式创制出纯粹的拼音式文字。

这种历史的契机在公元前18～16世纪之间出现了。说闪米特语的西亚游牧民族喜克索斯人逐步渗透到尼罗河流域，继而武装入侵了埃及北部并建立了十五和十六王朝。当其强大时，控制范围可能扩大到全埃及，前1580年左右才被逐出埃及。在这两个世纪中，喜克索斯人留给人类文明的伟大贡献莫过于字母了。闪语中"水"叫MEM，就借用埃及"水"的符号表示M。这样，他们从埃及象形文字音符中选用了22个符号，按照离素原则创造了最初的字母文字。1916年，英国埃及学家加德纳

(Gardiner)成功地释读了在西奈半岛发现的一些铭刻，这些西奈字母——也许就是喜克索斯字母，正是古埃及象形文字与腓尼基字母之间的中间环节。前 13 世纪左右，腓尼基产生了只表辅音而形体不同的两套字母系统，在毕布罗(Byblos)海港城出现了与西奈字母有承袭关系的 22 个字母，在乌加里特城出现了用楔形符号表示的 29 个字母。乌加里特字母渐渐淘汰，南北部通行由 22 个辅音所组成的毕布勒字母系统。

约公元前 1200 年，创造迈锡尼文明的希腊亚该亚人入侵克里特岛，曾在克里特线性文字基础之上改制成亚该亚音节文字。前 9 世纪，希腊文明的创造者多利亚人借用了腓尼基字母，通过增加元音字母形成元辅音字母表，也可能从亚该亚文字中吸取了个别符号。希腊字母表由厄脱鲁利亚人传给罗马人，前 7 世纪形成拉丁字母。前 8 世纪，闪米特人在腓尼基字母的基础上形成阿拉米字母。阿拉米字母东传印度等地，先后形成婆罗米字母和佉卢字母。

由此可见，拼音文字在借源方式中产生，世界上的字母文字都是借源文字。然而，借源方式不一定非出现拼音文字不可，比如阿卡德文、巴比伦文，就是非拼音式或意音式借源文字，这种情形在以汉字为本源的借源文字中表现得尤为明显。

## 四、以汉字为本源的借源文字

以汉字为本源，通过仿借原理和字型在东亚产生了多种借源文字，由此形成东亚汉字文化圈。

(一)仿照汉字体式创制文字的中国周边国家文字

1. 仿照汉字体式和吸收等韵原理而创制的朝鲜谚文。公元前朝鲜就接触到汉字，4 世纪末汉字已作为书面语通行。世传新罗神文王时期(681—692)，薛聪从汉译佛经对梵语音译中受

到启发而创造"吏读"。利用汉字的方法主要是训借和音借,训借类似日文的训读,而在标记专有名词与表示语法词缀时则采用音借。表示词缀的汉字比较固定且数量有限,逐步简化为偏旁或某些笔画,近似于日本的早期假名。1443 年,朝鲜总结了吏读的经验,吸收了汉语音韵学的成果及梵文、八思巴文的知识,创制了谚文即训民正音。谚文的字母笔画及拼写成方块,都是汉字体式的痕迹。正音文字问世后,并没有很快取代汉文和吏读的传统地位,直到 1895 年才将正音确定为官方正式文字,同时与汉字夹杂使用。1946 年北朝鲜停用汉字,朝鲜文才最终成为纯拼音文字。

2. 利用汉字个别符号而创制的日文假名。公元 3 世纪,汉字通过朝鲜传到日本,最初的日文全用汉字即等于把日语译成汉文,而人名只能用音译。5 世纪中叶,开始用汉字作为音符记录日本词语。8 世纪末编成的《万叶集》,除了个别汉字是训借,绝大多数汉字已经用作音符,被称为"万叶假名"。据说吉备真备以汉字正楷创制片假名;弘法大师空海以汉字草楷创制平假名。假名的创造者可能既懂得汉语音韵学,又懂得声明与胡书。理论上日语词都可以用假名来写,但在实践中由于日语中有大量的同音词,使得全用假名来书写成为不可能,因此日文不得不成为假名夹汉字的混杂文字。如果引进区别性符号分化日语同音词,即使汉字全部不用,日本文字也不是纯拼音文字。汉文假若要实现拼音化,因为同音词太多,势必也会陷入日本文字的这种境地。

3. 引进汉字并自制方块字的越南语字喃。1 世纪,汉字开始传入越南,后被称为"儒字"。13 世纪,越南人依据汉字的字型创造"字喃",从汉字六书中吸取了假借、会意、形声三种方法。17 世纪,西方天主教传教士用拉丁字母为越南语制定了字母

表，这以后越南开始使用字母文字，实现了文字类型的换用。

(二)仿照汉字体式创制文字的中国境内民族文字

1. 仿照汉字体式创制本族文字的南系文字。壮族借用汉字，起初采用音借、训借和全借(借词)，至少在唐代就已仿照汉字体式创制生字或方块壮字，后又有人采用形声法或反切法造字，夹在汉字中使用，但是方块壮字只是对记录壮语的汉字的一种补充，始终未能发展为壮民通用的文字系统。8世纪南诏时期，白族的民间就开始利用汉字记录白语。白文由借用汉字和增损汉字笔划自制的新字组成，现在民间仍有人使用。另外贵州的水家"反书"，把汉字反写，与汉字意义相同而字音不同。创制于唐代的彝文，由自制新字与借用汉字两部分组成，与汉字的渊源关系一方面表现在大量借用六国古文和后世真草作为彝文音节符号，另一方面表现在笔划部件与汉字具有共同性。

2. 仿照汉字体式创制本族文字的北系文字。契丹人在早期，曾由"汉人之陷番者，以隶书之半加减"，制定过一种契丹隶书。契丹神册五年(920)，契丹王朝制定大字；辽太祖三年(924)，契丹王朝又重新制定小字。契丹大字是表意类型的单体字，参照汉字仿造；契丹小字是表音类型叠体字，模仿字母制定。直到金章宗明昌二年(1191)才下诏罢用。西夏李元昊大庆年间(1036—1038)，党项人野利仁荣创制蕃文，这种后世称之为的西夏文，由汉字笔画组成，计有六千余字，其中会意、形声约占80%，同时有楷行草篆四体。1227年蒙古灭西夏，党项人逐渐同化于其它民族中，但西夏文于元朝中叶仍在甘肃宁夏一带流行，直至明代中叶个别地方还在使用。金太祖天辅三年(1119)，完颜希尹于"依汉人楷字，因契丹制度，合本国语制女真字"。制字基字借用了汉字和契丹字，以汉字为基字者大半减笔。熙宗天眷元年(1138)又颁布女真小字，现发现的与契丹大字相仿，

有意字也有音字。金亡(1234)以后,少数女真人还在使用,直到明朝中叶女真人弃女真文使用蒙古文,后女真满洲又仿蒙古文制订满文[1]。

南系文字的显著特点是,文字符号包括借用和自制两部分,借用汉字的方式通常是音借、训借和全借,自制新字的方式基本是变体(增损、反写)、形声和反切。北系文字的显著特点是,文字符号先后向仿造汉字和模仿字母方向发展,即使在该王朝灭亡以后民间仍然使用,直至被新的字母文字所代替。

## 五、东西方文字的不同是类型的不同

(一)自源文字不可能突变为拼音文字

自源文字或滞留在象形阶段、或使用音符和形声复合而进入成熟的意音阶段,但是都没有且不可能发展为拼音文字——"都没有"是文字发展史实,"不可能"是文化发展规律。某种自源文字,如果具备了音符法和形声法,记录语言的功能已经满足,而民族心理、文化传统又促使人们守住既有文字和历史文献不放。将既有的意音文字改变为拼音文字,意味着该民族的知识群体顿成文盲,意味着该民族的文化血脉突然割断。在历史上,只有家园沦丧、民族同化、语言湮没、文明中断,自源文字才会成为人类文明史中的化石。用后起的拼音文字取代传统的自源文字,不是自源文字的自然发展,而是自源文字的发展中断,不是这一民族的自行选择,而是其他民族的迫使就范。有人提出"楔形文字从苏美尔时代进入巴比伦时代和亚述时代,并实现了表形到表音的转变",与其说是波斯楔形文字(36 个音符、4 个意符和 1 个隔词符)发展到拼音阶段,不如说是波斯字母符号采

[1] 《中国民族古文字》,中国民族古文字研究会编印,1982 年。

用了楔形体式。字形和原理的二元模式并非波斯首创，腓尼基字母就有过一套楔形符号系统。

(二)拼音文字在借源方式中产生

古埃及人发明了圣书文字，在表意法之外又创造了表音法。由古埃及语的辅音词干而元音屈折的特点导致古埃及人创造出的离素原则，无疑是字母文字产生的重要因素。喜克索斯人借助离素原则和圣书符号草创字母，腓尼基人进一步发展和传播辅音字母表，希腊人完善成为元辅音字母表。随着民族文化接触，特别是宗教的传播，借源式拼音文字成为西方的主流。

(三)借用并非一定产生拼音文字

借源方式也可以产生意音类型文字。在西亚，由于苏美尔楔形文字没有产生离素原则，沿袭楔形文字的阿卡德文、巴比伦文、亚述文、赫梯文都是意音类型的借源文字。在东亚，由于华夏汉字没有产生离素原则，因此以汉字为本源的东亚借源文字也都是意音类型。通过音借和训借，形成直接借用式意音文字；通过借鉴形体和原理，形成间接借用式意音文字。由于语言类型不同，在汉语音韵学原理的影响下，日本和朝鲜的意音文字才向拼音文字方向转变。必须提醒，汉语音韵学是受梵文胡书的启迪形成的，而梵文胡书源于西亚腓尼基字母，因此东亚拼音类型借源文字的原理仍然导源于西亚腓尼基字母中包含的古埃及离素原则。

(四)文字发展基本规律的虚构性

文字发展的基本规律或盖尔波(Gelb)的文字进化论，其实就是把自源文字的产生发展与拼音类型借源文字的出现这两者捏合在一起，以单线进化论为基础虚构出来的一条规律。根据人类文字史，自源文字永远不会自发进化为拼音文字，而拼音式借源文字亦无须经历象形和意音阶段。

规律具有自发性、周遍性和不可抗拒性,但是无论是自源还是借源,从来没有一种文字的发展经历了象形——意音——拼音三阶段。自源文字有象形、音意阶段,无拼音阶段;而拼音式借源文字无须经历象形、音意阶段。作为文字发展的基本规律,假若与各别文字的演进历程不相符合,这一规律只能是虚构的。面对这样的历史过程,如果不准备作出"日本假名是汉字发展的拼音阶段"这样的判断,那么同样也不能推出"腓尼基字母是古埃及文发展的拼音阶段"、"波斯楔形字母是苏美尔文发展的拼音阶段"这样的结论。既然这个规律不过是错觉的虚构,为什么要作为汉文走拼音化道路的理论根据呢?

汉语语音的共起性导致了古代汉字假借原则的形成,埃及语语音的继起性导致了埃及圣书离素原则的出现,由此决定了东西方文字发展的不同趋势。意音文字和拼音文字是文字类型的不同,而不是发展阶段的不同。在人类文字史上,所谓"世界文字发展的共同的拼音方向"并不存在。反对"西化"者主张汉字拉丁化,也许自称是"洋为中用",但以拼音文字取代汉字,却无疑于"文化自杀"。关于汉字的前途,可以抛开虚构的"文字发展的基本规律"来重新考虑:深入探讨汉字与汉语、民族心理的相互关系;多方比较意音文字与拼音文字的异同;在信息处理中,不仅要促使汉字适应机器,更重要的是机器如何适应汉字。

## 第三节　语言类型与文字类型的制约关系

文字类型的传统分类是图画(或象形)类型、表意(或意音)类型与表音(或拼音)类型,由于图画类型以符号造型作为分类标准,与后两种类型的分类标准不相一致,有其不足,因此较新的分类法从符号功能出发,将世界文字划分为语段、表词、音节、

音位(辅音/元辅音)类型。像许多事物的分类一样,文字的分类同样呈现连续统状态,在这些纯粹类型之间必然存在一些混合类型,如辅音—表词型、音节—表词型、形声—表词型。根据符号造型,又可以将文字分为图画式、象形式、线条式等等。语言的形态分类法将语言分为孤立结构类型、粘着结构类型、屈折结构类型和复综(或编插)结构类型,同样这也是一个连续统。虽然闪含语言和印欧语言都属于屈折语,但是前者是以辅音构成词根而元音进行形态变化的屈折语,后者是元音、辅音既构成词根而又进行形态变化的屈折语,因此,可以把闪含语言称之为辅音词根型屈折语,把印欧语言称之为元辅音词根型屈折语。

传统的观点忽视了文字类型和语言类型的制约性,认为文字和语言之间是"靴子"和"脚"的关系。既然文字表现或记录语言,二者之间就不可能没有内在的制约关系。文字的特点或发展必然会受到若干非语言因素的影响,特别是政治需要的严重干预,但是从长远的观点看问题,最根本的制约仍然是语言自身的特点。因此,问题不在于"语言的特点并不能够制约文字的一切",而在于通过文字类型和语言类型的历史考察,揭示两者之间的制约关系到底如何。

## 一、自源文字类型的自在选择

### (一)文字画的综合与分解

原始文字由表达一定语言内容的文字画蜕变而来,文字画以综合方式与语段对应,或经分解方式与语词对应,将导致原始文字发展的最初两种方向。印地安、极北等民族的图画式语段文字,最重要的外部特征:它们不是一个个单独的图形,而是一组组或一套套彼此相联的图画,表达的是某个完整的、在图形上未分解为单个词的综合信息,与他们语言中的复综结构体相一

致。由于没有通常在别的语言类型中所见到的词单位,语言成份难以从复综结构体中分离出来,因此这些文字也就滞留于图画式语段文字阶段并充分发展。与之相反,在公元前第四千纪至第二千纪中期,苏美尔文字、古埃及文字及商代甲骨文都已经以表词文字的面貌出现。在独立使用单位为词的语言中,词的音义及其相应图符之间的直观联系促使图画的分解。可以推定,粘着型的苏美尔语、屈折型的埃及语和趋于孤立型的华夏语,这些语言中的象形符号一旦获得与单词的对应关系,便很快超越了图画式语段文字,而使表词方式在文字体系中占了主导地位,伴随着表音符号的出现,成为混合性表词文字。图画式的东巴文与复综语组图画式文字不同,纳西语的孤立型特点决定其文字是表词的,尽管残存语段文字的遗迹(如引导字),但它一字一音、一音一词,假借与形声都已出现,其自然发展趋向于形声—表词文字。任何一种文字画随着自身的发展,都有可能或多或少地出现表词字(至少有专有名词),然而是沿着图画式语段文字的方向使之充分发展,还是向表词文字方向过渡,显然与语言结构的特点有关。复综语中语言要素的综合占优势,非复综语中的语言要素的分解占优势,语言的综合导致了文字的综合,语言的分解导致了文字的分解。

(二)粘着型的苏美尔语与音节分析

文字分解为与表意词符以后,就有可能进一步将语词分解为音节单位。作为粘着型的苏美尔语,可以独用的语言单位包括词根与词缀。词根的语音结构类型有四种:V、CV、VC、CVC,单音节词根容易形成音节概念,而表达单音节词根的意符又易于转变为音节符号。苏美尔语广泛使用词缀来构形或构词,音节分解比较自然。这种语音结构的单一性与词缀的粘着性导致了音节符号的产生。通常是根词用意符,词缀用音节符号来表

达。单个辅音不担当词缀，因而没有单独的辅音符号。尽管音节符号使用频繁，但由于该语言中存在大量的同音词根，所以不仅不能放弃意符，而且必须运用部首符号以区别同音词，而促使苏美尔文字成为音节一表词类型。

（三）辅音词根屈折型的埃及语与辅音分析

从语言中析出辅音音素通常比析出音节要困难得多，但在某些语言里却变得较为容易。这些语言中构成词根是辅音，由屈折变化的元音表示不同的语法意义。闪含语言正是如此，词根通常由三个辅音构成（所谓三辅音原则），而夹杂在词根内的元音通过屈折构成派生词或作为不同的语法形式。反复出现的具有强调语义作用的辅音骨架，促使运用这种语言的人把辅音理解为至关紧要的语音要素。因为语法形式可以依据词序与语境来确定，因此文字只需要书写词根辅音也就记录了语义。由此可见，人类最初的辅音音素分析并非出于音素分析的自觉技能，而是这种语言的天然特征赋于闪含人这种天才，此盖语音学源于北非一西亚之内在原因。

古埃及语属闪含语言，约3300个根词中单辅音60、双辅音400、三辅音2200，四到六个辅音600。某些单辅音根词的意符用来做多辅音词的辅音符号；某些双辅音或三辅音的意符丢失了强辅音后的弱辅音而转变为单辅音符号。后来埃及语又向分析型方向转变，某些位置上脱落了阳性词尾辅音，这就增加了单辅音根词的数量，促进了辅音音符的发展。由于元音不以音符标明，致使辅音同音词根增多，不得不利用类符来加以区别而产生了形声字。因此，除了文化传统的力量之外，语言和文字的运用情况也决定了古埃及文字就是不可能演变成纯表音型，而必

须采取辅音—表词类型[1]。

(四)单音节孤立型的华夏语与形声原则

从商代甲骨文的造型可以推断它起源于原始文字画,但迄今在华夏地区未发现连环式的图画文字。刻划陶符与原始汉字不属于同一系统,而大汶口文化中发现的四个图画字是否与汉字具有渊源关系值得探讨。如果说华夏原始书写系统经历过图画形式,那极有可能是与东巴文相似的图画式表词文字。

甲骨文中绝大多数词语已经是单音节词,双音节组合甚少且多是专名。尽管推定殷商语可能是粘着类型,但粘着成分可能是音素,而汉语的孤立型倾向很早可能就发生了,周秦时期的汉语应已逐步孤立结构化。甲骨文造字以象形(包括在象形上添加符号的指事、多个象形复合的会意)为主,而常用形声字只有几十个。在用字中则广泛使用假借,可据形推其本义的不过百余字,而通过抽样检查,假借率高达75%[2]。由假借原则产生的音节字符与由音节原则产生的音节符号并不相同,假借中的音节字符为一个完整的能满足单词需要因而与词对应的语音单位。音节原则中的音节符号处于待拼状态,只与多音节词中的某一音节对应,这一音节不能满足一个词的语音表现形式。虽然苏美尔语中有许多单音节词根,但是与粘着性词缀合在一起就成为多音节词。华夏汉语一音一词,假借即以声托事,因声达义,而多音节词的语言中表达词缀的音节并不涉及语义层面。假借字符只有个别性、随机性,音节符号才具有普遍性,即凡此音节都可用此符号表示。音节符号的形成不但依赖单音节词的存在,而且依赖于多音节词系统的存在。单音节词为主的语言

① 以上参见B·A·伊斯特林《文字的产生和发展》。

② 参见陈炜湛《甲骨文简论》,上海古籍出版社1987年。

中，音节与词的对应关系，使之只能在有限范围内起表音作用。由假借字符演变而来的声符只能谐数字或数十字，不可能进一步演变为这个音系中的某一音节的普遍代表物。与之相反，处于多音节词网络中的单音节词，如苏美尔语或日语，一旦被用作表音符号，就包含着抽象为某一音节的普遍代表物的趋势。因此，单音节语中的假借字符与多音节语中的音节符号性质迥然不同。甲骨文虽然一字形一音节，但并不是文字类型意义上的音节文字，而是通过假借方式实现表音的单音节语词文字，即假借一表词文字。

假借方式的广泛使用，致使同音结构在字形上无从区别。文字区分度降低的结果，势必借助形旁补偿，于是形声原则逐步代替了假借原则。从甲骨文中的形声字初显端倪到秦汉篆文中饿形声字占 80%，经历了一千多年，汉字的类型就变成了形声一表词文字。

## 二、借源文字类型的自觉选择

自源文字由文字画演变而来，语言类型对文字发展的限定是自然而然的，故其选择为自在；借源文字则是通过文化接触引进的，尽管选择的历史契机或具体方式不尽相同，但一般总是表现为能动选择，故其选择为自觉。

### （一）不改变文字类型的因袭借用

1. 音节文字产生于粘着语，故非粘着语借用则不适应。操辅音词根屈折型闪米特语的巴比伦—亚述人，通过阿卡德文化借用了苏美尔文字，虽然广泛使用复合法，使后一个音节变成表辅音符号，但仍不适应巴比伦—亚述语。赫梯楔形文字借于巴比伦，每个字符只保留一个音节值，同时使用意符，使文字变得十分复杂，音节文字与元辅音词根屈折型的赫梯语也不相适应。

机械因袭借用方式，致使巴比伦—亚述文字、赫梯文字不能向适合自己语言的音位文字方向演进。因而通过向音节文字借用的方式产生不出音位文字，尽管借方的语言类型有这种要求。然而，有原始图画文字操粘着语的埃兰人，也通过阿卡德人借用了楔形音节文字，但他们根据埃兰语粘着型与音节结构简单的特点而使之向纯音节文字类型转变。

2. 辅音文字产生于辅音词根屈折语，故非辅音词根屈折语借用则不适应。元辅音词根屈折语借用辅音文字会使字母文字的类型发生变化，或如希腊另外创制元音字母而成为元辅音音素文字，或如印度使辅音符号元音化而成为音节文字。粘着语借用辅音文字，通常如维、蒙、满诸族仿希腊式而元音化。在辅音词根屈折语区域内，辅音文字也从最初的直接借用而出现了附加点线标明元音的方式，逐步演变为标元辅音文字。

东方字母一般都通过阿拉米文字而源于腓尼基文字。阿拉米语与腓尼基语相近，最初的文字几乎没有差异。亚述与巴比伦舍弃音节文字转而采用了阿拉米文字。希伯来人采用阿拉米文字，在字形上演变出方形、草书及斜体。巴利米拉文字是在叙利亚流行的阿拉米文字变体，由此形成福音文字，随着基督教的分裂而分化为三种文字。在阿拉米与巴利米拉文字基础上产生了表达阿拉米语的那巴泰文字，从三、四世纪起开始用来书写阿拉伯语。所有被阿拉伯人征服并信奉伊斯兰教的民族，阿拉伯文字都成其占统治地位的文字。在闪米特各国人民辗转使用辅音文字的同时，准确诵读祈祷文的需要使标元方法得以发展。用一定的辅音字母表示长元音的方法，称为“阅读之母”。另一种是用补充符号的方法，或如巴比伦文用缩小字母作为元音符号；或如希伯来文用行上下符号标示元音；阿拉伯文则用直线和曲线。它们已不是单纯的辅音文字，而成为标元辅音文字，这表

明辅音文字只是一种待补性的文字。

3. 形声文字产生于孤立语，故非孤立语借用则不适用。操粘着语的朝鲜、日本借用汉字，音借方式使汉字成为音符，导致吏读与假名的产生，最终创制出适合本族语言的音节文字。党项人制定以会意、形声为主的西夏文，白族用白文，壮族造生字，越南制字喃，多为操孤立语的民族借用形声文字。这些文字中，有些是随着民族的消亡而湮没，有些是由于西方文化的传播而换用了元辅音音位文字。

4. 元辅音音位文字产生于元辅音词根屈折语借用辅音文字的过程中，它适用于元辅音词根屈折语诸语言。欧洲的一些民族在利用拉丁字母表制作本民族文字时，一方面通常改变所借用字母表的结构和音值，改变的大小取决于拉丁语同新文字所书写的语言在语音结构上的差异大小；另一方面改变字母的形状甚至书写的性质，改变的大小取决于该民族书写传统及创制新文字时的文化风尚与政治状态。殖民扩张使元辅音文字在世界广泛流行，不但无文字语言以拉丁字母为基础创制文字，而且一些有文字语言也转而采用拉丁字母。这表明元辅音文字具有广泛的适应性。

毫无疑问，元辅音音位文字具有准确记录语音形式的功能，但对于那些同音结构较多的语言，文字不仅需要记录语音形式，而且必须显示同音结构在语序中难以辨清的语义差别。认为文字仅仅记录语音而不直接反映语义的观点，显然忽视了一个基本事实：最初的表词字是由文字画演变而来的，其视觉形式与语义直接相关，而听觉形式才与语音间接有关。过分强调文字记录语音形式的功能，势必导致对同音结构较多的语言也适用拼音文字的错误观念的产生。

(二)借鉴其它民族的文字体制而创制新文字类型的改制使用

1. 操辅音词根屈折语的腓尼基人借鉴辅音词根屈折语的埃及辅音一表词文字,创造辅音音位文字类型。作为新兴民族的腓尼基,没有传统文化的重负,具体历史机遇左右着其文字的选择。书写老腓尼基语初用毕布罗音节文字,继而使用原始伽南与原始西奈文字,以后又出现过乌加里特楔形辅音文字,公元前 12~10 世纪才产生了腓尼基字母。以上情况表明,腓尼基人选择埃及圣书符号与苏美尔音节原则相结合所产生的毕布罗音节文字不适合腓尼基语,选择根据埃及圣书符号和辅音原则创制的伽南与西奈文字,是埃及文字演变为腓尼基字母的中间环节,苏美尔楔形式样与埃及辅音原则相结合的乌加里特文字也不可能长久,最终还是依据埃及圣书符号与辅音原则所创制的辅音文字符合腓尼基语的基本特点。

2. 操元辅音根词屈折语的希腊人借鉴腓尼基辅音音位文字,创造元辅音音位文字类型。属于早期希腊的亚该亚人借用了米诺斯音节文字(克里特线形文字 B)。粘着型的米诺斯语以开音节为主,音素组合中不存在复辅音,音节容易分析出来,但音节文字不便表示希腊语词尾辅音与词首复辅音。后期的希腊多利亚人毁灭了克里特—迈锡尼文明后,亚该亚音节文字随之消失。与腓尼基人的贸易交往,给希腊人带来辅音文字。与腓尼基词根不同,希腊语的词根由元、辅音组成。许多语词要靠元音区分语义,只用辅音字母书写则字形相同。与腓尼基单一的词序也不同,希腊语因为具有丰富的形态变化而词序灵活,但仅采用辅音字母却不能暗示元音的屈折变化。辅音文字的元音化途径,在希腊为什么不像印度经由辅音音节化的方式,而通过创立特殊元音字母的方式呢?一方面音节文字对于具有多种形式

音节结构和复辅音及词尾辅音的希腊语极不方便；另一方面表达腓尼基喉塞音或半元音的一些字母被借用时，希腊人大概读成了希腊语的元音，也可能在读一些字母的腓尼基名称时，丢掉首辅音而读成了后面的元音。这样的读音在用以书写希腊语时固定下来，由此演化出希腊元音字母。希腊人除了增加了元音字母外，还根据希腊语的特殊语音结构改变了辅音字母的组成。元辅音类型文字的产生使辅音文字发生了根本改变，为语音的准确记录提供了便利工具。

3. 操元辅音词根屈折语的波斯人借鉴巴比伦—亚述楔形表词—音节文字，创制出楔形音节文字。公元前六世纪，西亚广泛流行巴比伦—亚述楔形表词—音节文字与阿拉米辅音文字。虽然这两种类型对波斯语都不合适，但迅速登上历史舞台的波斯人认为自己承袭了被征服的巴比伦王国的以往强盛，因此采取了具有西亚通用性质的楔形文字式样。音节文字不仅能表示辅音，也能表示元音，与辅音文字相比，相对适合元辅词根屈折语。同时在阿拉米音素原则的影响下，波斯人仅保留4个表词字，从而使波斯文字变成了准音节文字。

4. 操粘着语的日本人借用汉字，创制假名音节文字。日本起初通过训读与音读的方式，借用汉字表达日语。借助表词字可以表达不变化的词根，但记录表示语法形式的粘着性词缀却十分困难。始利用开音节汉字作为音节符号，继而演变出万叶假名，最终发展为平假名和片假名。由于日语中约有70%的词来自汉语，且有大量同音词，因此不得不采取假名夹汉字的书写体系。现代日语文章中，假名符号占总数60%，汉字占40%。除了在小学课本及电报中用纯假名音节文字，一般情况都采用的是表词—音节混合型文字。

（三）文字借用的复杂性

1. 选择的曲折性。与自源文字只有语言类型对文字类型的自然选择不同，借源文字则可以因历史机遇而自觉地多次选择，以求适应本族语的特点。早期的希腊亚该亚人借用米诺斯音节文字，但继起的希腊多利亚人并未采用。在文字空白两个世纪之后，又借用了腓尼基辅音文字，加以变革创造了适合希腊语的文字。巴比伦人因征服阿卡德借用了不适应自己语音的音节文字，后来又借用辅音文字，并用缩小的辅音字母作为元音符号以适合巴比伦语。信仰伊斯兰教的土耳其采用阿拉伯文字，因辅音文字不适合粘着型突厥语，20 世纪改用拉丁字母。

与以上这些历时多向选择有所不同，共时二向选择透露出较为复杂的文化心态。契丹于 920 年仿汉字制表词大字，4 年后又仿回鹘文造拼音小字。前者是崇尚中原文化，而后者是适应其语言特点。这种情形与古波斯既通行阿拉米文字，又采用具有文化传统的权威性楔形音节文字相似。腓尼基人的选择则更为曲折，经历了毕布罗音节文字、伽南西奈辅音文字、乌加里特楔形辅音文字多次选择以后，最终才采用了辅音文字。

2. 选择的多元性。与表达原则与符号样式都借自埃及文字系统的腓尼基文字的一元借用不同，还存在着文字选择的多元借用。乌加里特字母文字其表达原则借自埃及，而符号式样仿自苏美尔。蒙古文的字母构成借自回鹘文，而书写行款则是受汉文传统影响。朝鲜谚文的语音原理来自汉语等韵学和胡书蒙文，方块字式却仿汉字传统。另外，一些用拼音原理改造土著符号而创制的文字具有借用与自产的双重性。古日尔曼人向敌手拉丁人借来字母原理，使原先自创的线型词符具有音值而形成北欧文字。爱尔兰人熟悉了拉丁字母以后，也给自己的计算符号赋于音值而成为欧甘文字。

3. 选择的机遇性。文字的借用除受语言制约之外，往往还受诸多非语言因素的影响，使选择具有机遇性或偶然性。这些因素有地域交通、贸易往来、征服同化、宗教信仰、文化辐射、文化趋同、殖民引进和一语两国等。由于地域交通与贸易往来，印度商人熟悉西亚辅音文字，因此通过辅音文字元音化的方法产生了婆罗米文字，而由于地域遥远未能接受与之同语言类型的希腊式元辅音文字。喜克索斯人因征服埃及选用了辅音符号，阿卡德—巴比伦人征服苏美尔遂相借用楔形文字，都是征服者为被征服者的文化所同化。反之，阿拉伯、罗马帝国在其占领地区推行阿拉伯、罗马文字，则是被征服者受征服者的同化。前者是自觉同化，而后者被迫同化。阿尔泰语系的一些民族，由于信仰伊斯兰教而采用不适合粘着语的阿拉伯文字，是由于宗教信仰的相同。朝鲜、日本与越南借用汉字，党项、契丹、女真仿汉字制字，则受中原汉文化辐射所致。西班牙、葡萄牙、英、法等国的文字体制变成美洲、大洋洲中一些国家的文字体制，并在亚、非洲殖民地广泛流传，这是殖民引进。乌尔都语和印地语都源于古梵语，但由于宗教信仰和使用国家的不同，前者用阿拉伯文字书写，后者用梵文字母书写。生活在中国境内的柯尔克孜人与境外的吉尔斯人为同一种族，因信仰伊斯兰教长期使用阿拉伯文字，但是由于一语两国，后者 20 世纪先后改用拉丁字母和斯拉夫字母。一个民族先后借用多种文字，可能出于不同原因。回纥从 8 世纪采用突厥文；后又使用仿照中亚商业语言粟特文字母创制的回鹘文；10 世纪起因宗教信仰改用阿拉伯字母。如今中国境内的维吾尔人使用在察合台文基础上形成的现代维吾尔文，但境外的维吾尔人先后改用拉丁、斯拉夫字母，则为文化趋同。

## 三、语言类型对文字类型的刚性和弹性制约

(一)语言类型对自源文字发展导向的刚性限制

在人类文字的历史发展中,必须对自源文字的类型的自在选择与借源文字的类型的自觉选择加以区别。在自在选择中,语言类型对文字的选择限制似乎是盲目的,其实以不知不觉的方式体现选择的必然性,即何种语言类型只能自然发展出某种类型文字。尽管某些发展阶段对不同自源文字是共有的,但也不无差别。同样是图画式文字,复综型的印第安语为语段类型,孤立型的纳西语为表词类型。同样是表词文字,但粘着语促使其文字向音节方向发展,辅音根词屈折语促使其文字向音素化方向发展,孤立语促使其文字向形声化方向发展。

语言类型对自源文字发展导向的刚性制约表现为:1.图画式文字是否演变为表词文字,取决于该语言是否能析出词这样的语言单位;2.表词文字孕育出何种表音方式,取决于该语言易于析出何种语音要素。

(二)语言类型对借源文字类型选择的弹性限制

在借源文字的自觉选择中,因诸多因素的影响使语言类型的选择呈现偶然性,但在绝大多数场合语言类型总是通过改制或重选的方式表现出语言类型对文字类型的弹性制约。粘着语适应音节文字,粘着语如借用辅音文字会从使用标元法走向仿希腊式元辅音文字。元辅音词根屈折语借用音节文字,或将混合性音节文字改制成纯音节文字,或舍之而重新选择。辅音词根屈折语借用音节文字,只得弃之重选。辅音词根屈折语适应辅音文字。辅音词根屈折语借用辅音文字,会随着应用的要求而采取标元方式。元辅音词根屈折语借用辅音文字会通过两种方式实现元音化:或创制元音符号使之成为元辅音文字,或经辅

音元音化成为音节文字。单音节孤立语适应形声文字。单音节孤立语借用形声文字，会根据该民族语音系统创制方块字。粘着语借用形声文字：或通过“音借”创制音节文字；或根据拼音及等韵原理，创制合体音节文字；或采取双文字制度。元辅音音位文字具有准确记录语音的优良功能，所有的语言都能用它记录语音。复综语一般多引进元辅音文字，但切罗基语却采用了印地安人塞霍亚（Sequoyah），从自己的语感出发而设计的音节文字，使用160多年认为无必要做任何修改[1]。表达音位使复综结构太长，而音节文字使之相对缩短。

（三）汉字的表音方式难以摆脱语义羁绊

以单音节词为语根而孳生新词，在书写系统中或表现为相应的文字假借，这使汉字的表音方式一开始就注定难以摆脱语义的羁绊。一个语根可表示数个字根，但以某一字根为中心派生的一组同源字则共用一个声符，不仅依赖形旁指明义类，而且声旁在表音的同时也透视语义。同源借用使声旁不可能上升为纯音符，具有偶然性、随机性的非同源借用声符也只能在极小范围内使用。以单音节语素为主的汉语，使汉字声符不具备促成声符字母化的语言条件。

（四）表音文字的不表音基因

与表词文字既反映语音又反映语义不同，拼音文字排除了将反映语义，仅记录音节或音位。这对于多音节语、同音结构极少的语言比较适用。对单音节语、同音结构较多的语言，实行拼音文字则影响文字区分度。另外，音符表达功能的递减性或衰变性仍令人深为困扰。从根本上说，因为语音处于动态演变中，

[1] ［美］肖尼思·卡兹纳《世界的语言》（黄长著、林书武译），北京出版社1980年。

字母的表音准确性只是暂时的，表音文字类型中潜伏着不表音的"基因"。反之，意音文字类型却能够突破时空界限，因为语义比语音稳定得多。

（五）文字演进程序的两种含义

文字类型演变的进化论模式，是生物进化论在文字研究领域的机械搬用。文字进化论与文字发展的多元状态和语言类型与文字类型的制约关系格格不入。除了单纯的形体变化与符号增减以外，所谓文字的演进程序有两种含义：1. 某一个别文字体系内部不同表达方式的出现顺序。如汉字系统内象形、假借、形声方式的依次出现。2. 某一文字类型的演进程序。如由苏美尔音节—表词类型，经巴比伦—亚述表词—音节文字，发展为波斯音节类型；或如由埃及辅音—表词类型，经西奈、伽南文字发展为腓尼基辅音类型，又由此突变为希腊元辅音类型。前一种情况发生在自源文字内部，后一种情况是自源文字与借源文字的衔接发展。最先产生古老文明的民族创造了具有广泛影响的自源文字，但因其文化传统的保守性及语言方面的原因未能进行根本变革，尽管这些文字中已孕育出新类型的因素。只有那些没有传统文化的束缚而迅速走上历史舞台的新兴民族，才能将旧类型中的新因素解脱出来，创造适合自己语言的全新的文字类型。语言类型对文字类型的制约观，可以解释东西方文字类型发展的方向不同。

# 第二章　语言符号的任意性和可论证性

## 第一节　荀子的王者制名论与约定俗成说

### 一、荀子约定俗成说的通行曲解

中国语言学界的一个基本原则——名实关系的"约定俗成"，皆认为源于《荀子·正名》中的"名无固宜，约之以命，约定俗成谓之宜，异于约则谓之不宜"。今将对荀子"约定俗成"说的通行理解列出如次：

1.《辞海·语词分册》"约定俗成"条下引荀子说，解释为：

> 谓事物的名称是依据人们的共同意向而制定的，因而为人们承认和遵守。

2.岑麒祥《语言学史概要》(1958)中论述：

> (荀子)知道"名"是有社会性质的，……事物与名称之间本来并没有自然和必然的关系，用什么名称来指什么事物完全决定于社会的自由选择。这是语言学里面的一个基本原则，想不到这个道理早于二千多年以前已由荀子明白

指出。

3. 王力《中国语言学史》(1962)把荀子的这段话译释为:

> 事物的名称,无所谓合理不合理,只要人们共同约定就行了,约定俗成的是合理的,不合于约定的名称就是不合理的。

4. 邢公畹《谈荀子的"语言论"》[①] 阐释荀子这段话:

> 大意是说,词跟所代表的事物之间并没有固定的、必然的关系,社会约定了管这件事物叫个什么就叫个什么了。……这里值得重视的还不在于他能正确地解决在古希腊争辩了几百年的问题,而在于他指出语词成立的社会因素。"社会"这个术语,荀子称为"群",所谓"人之生不能无群"。关于"群"的理论,可以看他的《王制》、《富国》两篇。

几乎所有的现当代中国语言学著作中都作如是解,此不赘引。这种解释的权威性和公认性,使荀子的"约定俗成"说成了现代中国语言学中可以与索绪尔的符号任意性原则媲美的一条公理,姑且暂不讨论任意性原则的片面性与神秘色彩,然而,把荀子的"约定俗成"解释为"依据人们的共同意向而制定"、"决定于社会的自由选择"、"人们的共同约定",符合荀子《正名》篇中的本来思想吗?如果你阅读了荀子的文本,其答案只能是否定的。通行的"权威性"解释完全背离了《正名》篇中的原义。

① 载《人民日报》1962年8月16日。

## 二、荀子正名思想的王者制名论

在上引文章中，邢公畹指出："《正名篇》的写作目的在于阐明语言的社会实践，是直接为政治服务，而不是为了语言教学"，这是正确的，但是要对荀子的——而不是今人的——约定俗成说，有比较接近的——而不是背离作者原义的——理解，就不能不首先依据《正名》的上下文对荀子的"直接为政治服务"作一考究。

战国时代，百家争鸣，思想活跃，众说纷纭。荀子根据自己的学术思想，叹息"今圣王没，名守慢，奇辞起，名实乱，是非之形不明"，预言"若有王者起，必将有循于旧名，有作于新名"，坚持"王者之制名，名定而实辨"的政治主张。由此可见，在主张王者制名论的荀子看来，王者制名的事业分"有循于旧名，有作于新名"两个方面，一为因循先王所定之旧名，一为制定当朝后王之新名，在这里根本不存在什么"人们的共同意向"，"共同约定"，"社会约定"以及"社会的自由选择"等等。"制名"或"约定"这一过程的主导者是谁呢？从荀子《正名》的上下文中可以明白地看出：

> 后王之成名，刑名从商，爵名从周，文名从礼。
> 是散名之在人者也，是后王之成名也。
> 故知者为之分别制名以指实。
> （王者）然后随而命之……
> 此制名之枢要也。后王之成名，不可不察也。

"制名"或"约定"的主导者就是《正名》篇中反复出现的"后王"、"王者"及其任用的"知者"，根本无须舍近求远，到《正名》之

外的《王制》、《富国》中去寻找社会的"群"。与之恰恰相反，荀子对民间的擅作名者深恶痛绝。他不止一次地说，"故析辞擅作名以乱正名，使名疑惑，人多辨讼，则谓之大奸，其罪犹为符节度量之罪也"，"凡邪说辟言之离正道而擅作者，无不类于三惑者矣"。三惑者：惑于用名以乱名者，惑于用实以乱名者，惑于用名以乱实者。在荀子眼中，定名辨实，非民间所应为，当与符节度量一样，由王廷颁定。民间私自作名，"名"是邪说辟言，"人"要绳之以法。为了维护王者制名的权威，翦除"异端邪说"，荀子甚至要诉诸刑罚，这哪里有一点"社会自由选择"、"人们共同约定"的影子？退一步说，如果这种"制名"或"约定"有什么"社会性"，那么这种社会性，在荀子看来，大概只会是"王权性"，而绝不会是"民众性"。荀子的原话是：

> 王者之制名，名定而实辨，道行而志通，则慎率民而一焉。……故其民莫敢托为奇辞以乱正名；故其民悫，悫则易使，易使则公。其民莫敢托为奇辞以乱正名，故壹于道法而谨于循令矣。如是则迹长矣。迹长功成，治之极也。是谨于守名约之功也。

使用语言的芸芸众生只能俯首听命，奉守王法，而不应有定名辨实的权利。荀子所提倡的王者制名论的实质，就是为了使其民不作奇辞，听从使唤，"谨于守名约"而达天下大治。

## 三、荀子正名思想的约定俗成说

明确了《正名》篇中荀子的王者制名论的主导思想，再来揭示"约定俗成说"的含义就比较客观，就比较有可能接近作者的用心，而不至于郢书燕说。为了有助于理解和说明问题，不妨把

本文开头所引的《荀子·正名》中的那句话再多引一些：

> 名无固宜，约之以命，约定俗成谓之宜，异于约则谓之不宜。名无固实，约之以命实，约定俗成谓之实名。名无固善，径易而不拂，谓之善名。

约，定也；定，约也。“约定”是同义并列结构，意即规定、确定、制定。俗，《说文》：习也；段注：习者，数飞也；引伸之凡相效谓之习。因此“俗”亦有相效之义。又《汉书·地理志》曰：(凡民)好恶取舍，动静无常，随君上之情欲，谓之俗。“俗”又有随君上之情欲之义。成，成就，实行，推广也。“俗成”亦是并列结构，即相效实行。那么，现在要问——“约定”与“俗成”，是谁来“约定”，谁来“俗成”？是人们“约定”、人们“俗成”吗？不是。根据上一节的讨论和荀子这段话的上下文（上文为“随而命之”，下文有“此制名之枢要也。后王之成名，不可不察也”），显而易见，只能是王者约定，王者使其民相效推行。

因此，荀子这段话就是说：

> 名（相当于索绪尔的能指）没有固然的适宜，王者制定事物之名，王者约定而民众相效的名，就是适宜的名，与王者制定不同就是不适宜的名。
>
> 名没有固定的实（相当于索绪尔的所指），王者确定名的含义，王者约定而民众相效的名，才是具有实义的名。
>
> 名无所谓天然的妥善，只要明白无歧义，又不违背实义，这样的名就可谓妥善之名。

荀子的制名或名约思想，在这段话的意思为三层：1.“正名”；

2."辨实";3."名副其实"。荀子主张,不论是有循旧名还是有作新名,不论是因实定名还是据名辨实,都是王者智人之事。名的制定权、实的解释权,必须牢牢把握在统治者手中,其民除了相效以外,别无权利干涉。荀子是说"王者制名,其民俗成(相效)",而不是"人们(或社会)的共同意向决定",荀子是说王权规定、强行推广、翦除异说、维护正统,而不是"社会自由地选择"。作为先秦"名实之争"中的"王权规定论者",荀子从未也不可能"正确地解决在古希腊争辩了几百年的问题"。

人们都习惯于说"社会约定",但是这一约定的具体历史过程,似乎没有人详细描写过。因此,这种浑浑沌沌的"社会约定"使人们只能作恍恍惚惚的理解。如果撇开荀子这段话上面笼罩的政治色彩,那么可以看到荀子实际上粗略地描写了这个过程,它似乎可以表述为:字词的初创在个人(并非独立于某个语言社团之外的个人),使用范围的不断扩大(不排除行政强权的推行),导致这个字词为较大的语言集团,直至全民所接受。荀子的"政治语言学",其特色就在于把个人在语言活动中的初创性与集团在语言活动中的认可性紧紧地结合起来了。

## 四、荀子约定俗成说曲解的溯源

哲学史、逻辑学史、思想史方面的论著,在论述到荀子思想时,都要涉及到"约定俗成说"。下面就几本具有代表性的论著考察一番。

冯友兰《中国哲学史新编》(1965)在"荀子的逻辑思想"这一节里,有两处解释"约定俗成",现引出来做一比较:

什么"名"指什么实,本来是没有一定的。荀子说:"名无固实,约之以命实,约定俗成,谓之实名"。例如,马本来

> 也可以叫作狗，狗本来也可以叫做马。可是在制名的时候，大家有个共同的理解，马叫马，狗叫狗；后来大家都是这样叫，成为习惯。……荀子认为什么“实”用什么“名”，是由于“约定俗成”。荀子的这个思想，表示他认为“名”是社会的产物，是具有社会性的。（537 页）
>
> 荀子认为名是由“王”制定的。在开始制定的时候，甚么“实”用甚么“名”本来没确定，但是既经“约定”而且“俗成”之后，那就成为“正名”，不能随便改动。……他也确切地认识到名是表达思想的必要工具。他要把制名之权归之于“王”，使统治者能统一思想，只提倡他所需要的思想。
>
> （541 页）

两处阐述，哪一处较符合荀子的原意呢？是 537 页的“在制名的时候大家已有个共同的理解”、“名是社会的产物，是具有社会性的”，还是 541 页的“要把制名之权归之于王，使统治者能统一思想，只提倡他所需要的思想”呢？显而易见，符合荀子原意的是后面一处的“王者制名论”。需要特别提醒的事，“大家约定”与“王者制名”是两个完全不同的概念。如果从荀子的约定俗成说中断言荀子主张“名”是社会的产物，因而“名”具有社会性，那么就得进一步追究，这种“社会性”是什么具体含义呢？充其量不过是王者制名以责其民相效，而要达到这种普遍接受程度，可能就离不开强权与暴力。虽然，并不打算把王者智人以至整个统治阶层排斥在社会之外，但是在制名的“王权性”和一般的“社会性”两者之间划等号，不觉得概念的含义变了吗？

胡适在博士论文《先秦名学史》[1] 中说：

[1] 《先秦名学史》（李建钊等译），学林出版社 1983 年。引文参照了英文版。

荀子否认名的神秘起源,代之以感觉经验和理智活动产生名的这种理论。然而,他保留这样的观点,即名首先是运用政权的法令"制定"的。虽然他并不否认以后的政府有同样的权力来制定新名,批准和改正那些未经政府批准而时常出现的名。荀子提出的观点与那种维护名的原始的标准含义的旧观点不同,他认为凡经社会约定(through social conventions),或者政府批准而变成流行的名就是正确的。政府应当批准的名,是那些因某种默契而已经流行的名。一切名词术语的创新都要由法律禁止。

胡适又进而对荀子的"名无固宜……约定俗成谓之实名"那段话加以阐发道:

这些原则赋予儒家的正名学说以新的意义。这些原则认可名的社会来源(the social origin of names),因而摒弃了儒家"象"的理论,也摒弃了那种认为名总是要根据它的原始的标准含义来理解的观点。依照新的见解,在交际中被采纳的含义才是正确的含义。名的正当性依据法令,圣王通过法令来审定那些通行的名。在经过法令批准以后,凡是企图创造新名的,就和私造官印、官称及官尺一样,是犯罪的。

胡适的阐述大致上符合荀子思想,不过行文中已出现一些主观的发挥。他提到的"经社会约定"和"名的社会来源",里面已隐含着"社会约定俗成"的萌芽。只要一念之差,再略加发挥,"王者约定、其民俗成"就会变成"大家约定俗成"了。回国以后,胡

适在北大讲授哲学史，在教材《中国哲学史大纲》(上卷，第十三版)中，这一主观的发挥便过了头。胡氏说：

> 荀子讲正名只是要把社会上已经通行的名，用国家法令制定；制定之后，不得更改。（329 页）
>
> 此处当注意荀子知道名有社会的性质，所以说“约定俗成谓之宜”。正名的事业，不过是用法令的权力去维持那些“约定俗成”的名罢了。（335 页）
>
> 名约即是“约定俗成谓之宜”，荀子的意思是要问大家的意见如何……（338 页）

经过这样的发挥，《正名》篇的王者有循于旧名、有作于新名的两方面，在“是要”、“不过”这些字眼的限定下只剩下对旧名的审定。旧名的“先王性”(刑名从商，爵名从周，文名从礼）被“社会上已经通行的名”的说法所掩盖；含混的“名有社会的性质”取代了“名”之上的王权统治烙印；而在将“约定俗成谓之宜”通俗化为“要问大家的意见如何”的同时，也就背离了荀子的“(王者）约定(民众)俗成”，由此流变成“依据人们的共同意向而制定”，“什么名称来指什么事物完全决定于社会的自由选择”。

以己强人，以今臆古，断章取义，主观发挥以致曲解荀子“约定俗成”的原意，胡适是始作俑者，其后的学人乃道听途说、三人成虎，语言学界大概是为了与索绪尔符号任意性原则媲美而一时冲动的盲目“移植”罢了。由此可见，曲解荀子的“约定俗成”，确实为“胡说”。

## 第二节　索绪尔符号任意性原则的失误与复归

### 一、对符号任意性原则的质疑

费尔迪南·德·索绪尔(Ferdinand de Saussure, 1857—1913)是本世纪最著名、影响最深远的语言学家之一。他的学生根据课堂笔记和他的一些手稿及其他材料整理编辑成《普通语言学教程》[①]。在此书中,索绪尔提出一个著名的论断——"语言符号是任意的"(102页),符号的任意性原则成为索绪尔语言符号理论的基石之一。

从那以后,国外的大多数有着索绪尔倾向的著作,都以不同的方式论述与之相关的基本问题,然而,同时有些学者(包括他的学生)却试图发展或修正索绪尔的语言符号任意性学说,甚至对此作了批判。1939年,本维尼斯特(E. Benveniste)证明符号根本没有日内瓦学者所设想的那种任意性;确切地说,符号对外部世界来说是任意的,但在语言中它却不可避免要受约束。1940年,沙·巴利(C. Bally)——索绪尔的学生和《教程》的主要编者——以索绪尔提出的关于绝对任意性符号和相对任意性符号之间的差别为出发点,区别出外部形式上的约定性和内部关系中的任意性。并由此得出结论,符号的生命就存在于以内在必需的联想为基础的全约定和以外部随意的联想为基础的全任意这两极之间。语言的符号性质问题,在前苏联语言学界长期以来曾经是禁忌语,直到50年代才打破禁区。1954年,布达哥

① 《普通语言学教程》(高名凯译),商务印书馆1982年。以下简称《教程》。引文所注页码即中译本页码。

夫(P. A. Будагов)说:"我们认为,'语言符号'一般只限于在其低级形式(音、部分词素)中才是不可论证的,但在其高级表现形式(词)中,'语言符号'总是趋向于可论证性的。"1955 年,斯米尔尼茨基(A. И. Смирницкий)写道:"不可论证性(任意性)适用于简单的、不可分解的或者相当孤立的、溶为一体的凝固性单位。至于在复杂的结构中起作用的则已经是可论证性的原则了。当然,它是跟第一个原则同时起作用的,因为复杂的结构中包括简单的单位。"① 这二位提出的语言中任意性成分和可论证性成分相结合的主张,与巴利的观点相近,而巴利的观点又脱胎于索绪尔的绝对任意性和相对任意性的区别。与本维尼斯特的观点不同,这些观点只是进一步强调相对任意性。1962 年,兹维金采夫指出:语言符号的任意性或不可论证性是明显站不住脚的论断,语言符号的任意性原则本身对语言本质来说不是不可缺少,而是在历史上可有可无的东西。

以往的中国语言学界,几乎不假思索地就全盘接受了任意性原则。在所有语言学文献中,只要涉及到语言符号的性质,都一无例外地重复索绪尔的这一论断,充其量把荀子的"王者约定,民众俗成"曲解为"社会或人们约定"。殊不知概念与音响形象的联系,索绪尔起初是用惠特尼的约定论来回答,即认为只有约定可以把不同的声音和不同的意义结合成一个统一体。但后来他认为,无论是声音还是意义,在结合成符号以前,两者都是不定形的。因此,索绪尔最终放弃了"约定论",转而提出了"任意性"②。

① 转引自[苏俄] B·A·兹维金采夫《普通语言学纲要》(伍铁平等译),商务印书馆 1981 年。

② 徐志民《欧美语言学简史》,学林出版社 1990 年。

## 二、任意性原则证明过程中的三个失误

先就《教程》第一编第一章§2《第一个原则:符号的任意性》的内容提出批评。

索绪尔认为:语言符号是一种两面的心理实体,一方面是概念,另一方面是音响形象。"语言符号连结的不是事物和名称,而是概念和音响形象。"(101页)其实,语言符号首先是事物与名称(先秦哲学叫"名实"之争,传统小学叫"名物"训诂。中国古代的"名"相当于现代汉语的"词")的关系。语言中的名物符号产生于能动的主体感知具体事物并给事物赋予名称的认知活动之中,只有再把语言符号或语词加以解析才有音响部分与概念部分。语言符号联系的是事物(实在的或虚幻的,后者不过是前者的变形)和名称,这个事物可以化为概念移入人脑,但事物≠概念,这个名称可以变成音响形象潜入心灵,但名称≠音响形象。索绪尔从社会心理学派出发,才提出语言符号的任意性原则,在他看来:"能指和所指的联系是任意的,或者,因为我们所说的符号是指能指和所指相联结所产生的整体,我们可以更简单地说:'语言符号是任意的'"。(102页)可见,符号的任意性是指存在于能指和所指之间的某种联系。他是如何举例证明自己的原则呢?"'妹妹'的观念在法语里同用来做它的能指的s-ö-r这串声音没有任何内在的联系;它也可以用任何别的声音来表示。语言间的差别和不同语言的存在就是证明:'牛'这个所指的能指在国界的一边是b-ö-f,另一边却是o-k-s。"(103页)

索绪尔在同一语言系统内考察,断言"妹妹"的观念在法语里也可以用别的任何声音来表示。用什么声音呢?他不可能并且也讲不出来。法语中现实的具有历史继承性与系统性的音义结合关系,被他撇开而代之以玄虚的"也可以"。面对所指和能

指的结合是历史地形成的并处于某一系统中这一状况，要弄清“妹妹”和 sor 这串音节之间的结合关系就必须溯源。索绪尔认为现代法语中的 berger“牧童”是绝对任意的，但如果溯源寻踪，berger 是由古代法语 bergier 变来，而 bergier 是由民间拉丁语 berbicariu(m)变来的。berbicariu(m)在拉丁语中是 berbicarius 即“牧羊人”，与 berbix“绵羊”有关。(185 页)显然，对所指和能指的关系不加历史的探讨，而以“任意性”一言蔽之，是任意性论证中的第一个失误。

接着，索绪尔又列举出不同语言系统之间能指和所指结合的差别，企图证明同一语言系统内能指和所指结合关系的任意性。我们不必坚持法语和德语同属印欧语系，像“牛”这样的基本词，尽管两种语言现在的叫法不同，然而可能出于同一祖语。也不想举出汉语中“啊[a]”、英语中的“ah[a:]”和俄语中的“a[a]”都表示感叹来作为反证，以说明不同系统之间亦有相同。而是要指出，索绪尔的这种论证方法是不合适的，不同语言系统之间的差别，只能在各自系统内部溯源论证。以不同语言系统之间能指和所指结合关系的差别来证明同一语言系统之内能指和所指结合关系的任意性，是索绪尔任意性论证中的第二个失误。诚然，同一语言系统内某些所指可以有不止一个能指即“一物多名”或“异名同实”的现象。但是，它们各有自己的来源，它们之间有古今之别、方雅之别，可以分头溯源。归根结底，仍然是系统的差别(历时系统与区域系统)。这种一个所指有不止一个能指的现象，可以称之为能指或命名的可选择性，与符号的任意性是两个完全不同的概念。

索绪尔在否定能指是象征这一观点时说：“象征的特点是：它永远不是完全任意的；它不是空洞的，它在能指和所指之间有一种自然联系的根基。”(104 页)据此可知，索绪尔所讲的所指

和能指之间的那种“内在的关系”（103 页），就是指“自然联系的根基”。换而言之，有自然联系的就不是任意的，反之，任意性就是无自然联系。他进一步定义：“任意性这个词还要加上一个注解，它不应该使人想起能指完全取决于说话者的自由选择。……我们的意思是说，它是不可论证的，即对现实中跟它没有任何自然联系的所指来说是任意的。”（104 页）由此可知，索绪尔的任意性 = 没有内在的关系 = 没有自然联系 = 不可论证。

在讲到“名”“实”（相当于能指和所指）关系时，坚持声音和意义之间不存在任何自然联系这一观点的人们，常爱引用马克思的一句话：“物的名称，对于物的性质，完全是外在的”[1]，从而断言“凡企图寻找事物名称和事物性质之间的关系的人，都不可避免地陷入了唯心主义的泥坑”[2]。马克思的意思是：物的名称不是物的自然属性，而是物的自然属性之外的东西，是人所赋予物的。虽然谁也不会说物的名称等同物的自然属性，但马克思并没有断言，物的名称与物的自然属性之间不存在某种自然联系。因此，名称是否是物的自然属性是一回事，名称与物之间是否存在着某种显性或隐性的自然联系又是一回事。企图用马克思的这段话作为名实之间不存在自然联系这一判断的权威性佐证，实在是文不对题。

中国古代有些学者主张“名物各有来源”。汉代刘歆《论方言书》中认为：一切命名应当皆有证验，“非徒无主而生是也”。这种寻求名物来源的探讨从周秦时代即已开始，至东汉刘熙《释名》总成，其方式固然质朴而易失之于疏，但透过这些次生文化层可以窥见原生文化层中命名的端倪或元语言要素。18 世纪

①　马克思《资本论》，89 页，人民出版社 1953 年。

②　王力《中国语言学史》，山西人民出版社 1981 年。

德国的赫尔德(J. G. Herder)在《语言起源论》中认为，人类能从许多冲击他心灵深处的、各种宛如汪洋大海的感觉中挑选出一朵浪花来，紧紧抓住它，把它转成语词。一个人看见一只羊，依他的感觉判断，羊是白色的、多毛的和驯良的，但是留给人们最深的印象是叫声"咩咩"，人们就把这叫声看作羊的区别性特征，于是"咩咩"的叫声就成了羊的名称。显而易见，这两者之间存在自然联系的根基，因而是可以论证的，并非任意的。当然，处在另一语言系统中的人可以从另一角度，抓住羊的另一特征来命名，但是这种不同语言系统从不同角度给一物命名各异的情况，并不能作为否认以羊的叫声给羊命名这二者之间有自然联系的证据。

一个所指可以从不同角度获得几个能指的现象，在上文中曾用命名的可选择性这个术语表述。这里要加以补充说明，命名可以不同，但对象仅限于一个，你只能针对这一对象作"自由"的选择，因而也就是"不自由"的。命名必须接受对象本身性质的制约，这就是命名的限制性。限制命名的因素是多种的，对象的性质只是其中之一，然而是最基本的。可选性与限制性构成了命名过程中的一对矛盾，使命名活动呈现出生动多彩的局面。如家用有孔煤球，北方多蜂窝则称之为蜂窝煤，湖南出藕则称之为藕煤，南京人则直呼煤基。

对于这种有自然联系的拟声词，索绪尔则认为："人们可能以拟声词为依据认为能指的选择并不都是任意的。但拟声词从来不是语言系统的有机成分，而且它们的数量比人们所设想的少得多。"(104 页)在这里，试图用共时的比例掩盖历时的起源。就现代词库而言，象声词的比例极小，这是因为：1. 存在大量的合成词；2. 有些象声词在历时的发展中丧失了原初的特征。但是，通过象声——这一最古老、最自然、最基本的命名方式——

窥见了人类自然语言中符号的奥秘。现代汉字中象形字比例极小且多已变得非象形,但是又有谁据此否认汉字起源于象形呢?莱布尼兹(Leibniz)在《人类悟性新论》中把人类一般语言机能的起源称为"原始语言"的起源,把各个个别具体语言的起源称为"派生语言"的起源。他认为前者是根据摹声原则创造的,后者在更多的情况下是根据符号原则发展出来的。叶斯泊森在《语言的本质、发展与起源》中把有关语言起源的理论归纳成为四个学说:摹声说、感叹说、声象说、喘息说。通过这些方式产生的能指,与所指之间的自然联系是明显的。

以象声方式创造的语词不仅以原始符号的面貌隶属语言系统的有机成分,而且数量也比索绪尔所认定的多得多。齐佩瑢认为:自然界的声音可分成物体自发的声音和物体客观存在外力而发的声音两大类。今姑以《诗三百篇》为例:关关雎鸠、雍雍鸣雁、鸡鸣喈喈、鸟鸣嘤嘤、奏鼓简简,伐鼓渊渊……[①]。关于《诗经》象声词的研究表明,象声词是形容词的一个组成部分[②]。用象声方式创造的词,在汉语中的名词、动词、形容词、副词、叹词等等,可以统称之为象声词。

索绪尔不仅否认拟声词是语言的有机成分,臆断其数量极少,而且还认为:"真正的拟声词,不仅为数甚少,而且它们的选择在某种程度上已经就是任意的,因为它们只是某些声音近似的,而且有一半已经是约定俗成的模仿。此外,它们一旦被引进语言,就或多或少要卷入其他词所经受的语音演变、形态演变等等的旋涡;这显然可以证明,它们已经丧失了它们原有的某些特

① 齐佩瑢《训诂学概论》,中华书局1984年再版。

② 陈波《从诗经看先秦形声词的语法特点》(油印本),华南师范大学1985年硕士论文。

征，披上了一般语言符号的不可论证性的特征。"(105 页)模仿总是近似的，这毕竟要受发音器官与发音习惯的限制。但是，近似的模仿与任意命名不同，透过模仿的近似性可以看到模仿声响与模仿物之间的自然联系。象声词既受语音系统的制约，历时演变又使它丧失了初生时与模仿物之间自然联系的某些特性，但并不等于原初就从未有过自然联系的根基。把符号的历时演变性与符号的不可论证性混为一谈，是索绪尔任意性论证中的第三个失误。

从根本上说，语言符号的论证性问题是个历时语言学问题，索绪尔总是以共时语言学的截面研究来混淆或掩盖语言符号可论证性的历时追溯。他也感觉到符号任意性原则无法一以贯之，故退而言之："拟声词和感叹词都是次要的，认为它们源出于象征，有一部分是可以争论的。"(105 页)然而他又要完善自己的原则，以致于不得不在方法上用不同语言系统之间的差别性来证明同一语言系统之内的任意性，用共时的比例掩盖历时的溯源，在模仿的近似性与选择的任意性两者之间划等号，把符号的历史演变性和不可论证性混为一谈。

## 三、绝对相对任意性划分中的三步复归

索绪尔在《教程》第二编第六章《绝对任意性和相对任意性》中写道："符号的任意性的基本原则并不妨碍我们在每种语言中把根本任意的，即不能论证的，同相对任意的区别开来。只有一部分是绝对任意的；别的符号中却有一种现象可以使我们看到任意性虽不能取消，却有程度的差别：符号可能是相对地可以论证的。"(181 页)

语言事实能纠正研究的偏见，而思维定势却使人不知不觉地偷换概念。把任意性区分为绝对与相对，又说明相对任意＝

相对地可以论证,从而把可论证性纳入不可论证性。羞羞答答地使相对任意性改头换面,是索绪尔的任意性原则向可论证性原则复归的第一步。

那么,他的可论证性对象是哪些呢?"法语的 vingt'二十'是不能论证的,而 dix-neuf'十九'却不是在同等程度上不能论证,因为它会使人想起它赖以构成的要素和其他跟它有联系的要素。"(108 页)诸如此类,例子都是一些复合词或派生词,用他的术语说:就是由两个或几个连续的单位组成的,因而可以分解为次单位的组合。他进而阐明:"句段的分析越是容易,次单位的意义越是明显,那么论证性就总越是完备。"(183 页)索绪尔把可论证性对象限制在这些合成结构中,提出可以通过分解来论证。只要把合成词与单纯词的数量做一比较,显而易见,即使根据索绪尔的两种任意性理论,相对任意性或可论证性的对象也是占大多数,况且单纯词中也有许多是由派生而来。

因此,索绪尔接下来关于论证性永远不是绝对的阐述也就显得毫无意义。他以为即使在最有利的情况下,论证性也永远不是绝对的。这不仅因为可以论证的符号的各个要素本身是任意的,而且因为整个要素的价值永远不等于各部分的价值的总和。(183 页)各个要素本身是否是任意的,与单纯结构是否可以论证有关。至于整个要素的价值永远不等于各部分的价值的总和,这一论断从词义学角度来讲是合适的,但与构词法和词源学却相去甚远。

索绪尔是这样表述相对可论证性的含义:"1.把某一要素加以分析,从而得出一种句段关系;2.唤起一个或几个别的要素,从而得出一种联系关系。"(183 页)请注意,在《符号的任意性》中,索绪尔的任意性=没有自然的或内在的联系,反之,可论证性是说能指和所指之间有自然联系的根基。而在此处,可论证

性的含义变成了——句段关系的可切分与联想关系的可类比。换而言之，就是把合成结构分解为次单位、通过次单位的含义来说明合成结构的可论证性。根据索绪尔分别规定的可论证性的含义，可将其一分为二：通过自然联系，即名与物之间的关系而直接考察进而阐释者，称之为直接可论证性；通过切分类比，以次单位的含义对合成结构加以阐释者，称之为间接可论证性。

在判别某一对象是否具有间接可论证性时，要考虑到历时的因素，因为对能指和所指之间的论证，实质上是一种历史溯源性研究。一方面要考察句段关系是否磨灭。有些词从现代语言的结构来看是单纯性的，而在历史上却是合成性的。如德语sterm（星星）、horn（兽角）；如法语 forge（铁匠铺）、berger（牧童）[1]。扬雄《方言》曰："美状为窕，美心为窈"，刘熙《释名》曰："匍，捕也，藉索可执取之言也；匐，伏也，伏在行走也"。现代汉语中认为联绵词的此类，在古代汉语中却析为复合词。另一方面，要考察联想关系是否丧失。"蝦/鰕（虾）"之得名之由或可论证性，用现代眼光孤立地看就茫然无识。而汇集了同族的瑕（玉小赤）、霞（赤云气）、葭（红芦苇）、騢（赤白杂毛马）、赮（赤色）等，从中提取出共同义素"赤"后，就可从[＊hea·赤色]这一语根中，推断出"虾"之得名缘于其色"赤"[2]。

由于索绪尔立足任意性，因此他对整个语言系统的概貌只能作如下的描述：整个语言系统都是以符号任意性的不合理原则为基础的。这个原则漫无限制地加以应用，结果将会弄得非常复杂；但是人们的心理给一大堆符号的某些部分带来一种秩序和规律性的原则，这就是相对论证性的作用。"语言机构"只

① 《普通语言学教程》中译本 185 页校注。

② 张永言《词汇学概论》，华中工学院出版社 1982 年。

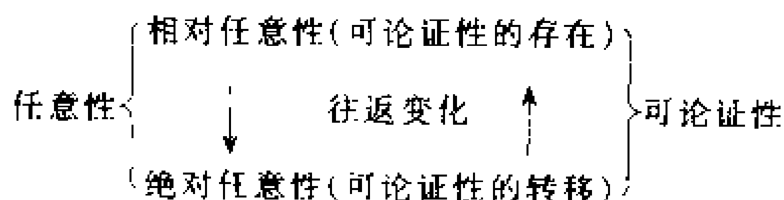
是一个本来就很混乱的系统作局部的纠正的研究对象。（184页）果真如此，人们的心理只能给某些部分带来规律性的原则，而另一部分中的不合理原则是什么东西带来的呢？是谁在冥冥之中制造系统的混乱而与人们的局部纠正抗衡呢？索绪尔的描述会不会把人们导向神秘主义呢？

与索绪尔的论断相反，事实上不能论证的语言是不存在的；一切可以论证的语言，在理论上是存在的。由于语言的历时发展、某些符号可论证性的磨灭，导致各种语言都按不同比例包含两类要素。索绪尔依据“不可论证性达到最高点的语言是比较着重于词汇的，降到最低点的语言是比较着重语法的”（184页）这一前提，断言“超等词汇的典型是汉语”，因而汉语是不可论证性达到最高点的语言。虽然上古汉语以单纯词为主，但周秦时代就有合成词，汉唐以降日趋增多，除了这种复合构词，汉语中单纯词派生孳乳的情况丰富多彩。正是索绪尔认为的这种最不可论证的语言，在公元二世纪就出现了探讨名源、“论述指归”的专著《释名》，提出了“名之于实，各有义类”的论断。这种源头可追溯至公元前的学说，从汉代声训经宋代王子韶“其类在左，其义在右”，至清代王念孙诸人“引申触类，不限形体”，古代音义家一直在探究汉语符号的可论证性。章炳麟的《文始》、高本汉的《汉语的词族》、藤堂明保的《汉语语源词典》、王力的《同源词典》也都在以明语源、聚同源的方式进行汉语符号可论证性的探索。

事实上，绝对任意性也并非不可论证。索绪尔举出一些法语例子，许多拉丁语中可以论证的符号，到了现代法语中都不能论证了。索绪尔一方面客观地认为“在数以百计的例子里都可以看到这种转移”，另一方面又固执地断言“已恢复到作为语言符号主要条件的绝对任意性”（185页）。显而易见，现代法语符号的可论证性自然当根植于拉丁语。以现代的论证性丧失来否

定历史上论证性的存在，用“恢复”一词来描写论证性的磨灭，则显得本末倒置。但是在此处，索绪尔对绝对任意性又有了新的解释——可论证性的转移。如果说绝对任意性是可论证性的丧失，那么索绪尔的任意性原则又向可论证性原则回归了第二步。

在索绪尔的笔下，语言机构是这样运动的：“在一种语言内部，整个演化运动的标志可能就是不断地由论证性过渡到任意性和由任意性过渡到论证性；这种往返变化的结果往往会使这两类符号的比例发生很大的变动。”（185 页）这种往返变化可以图示如下：




任意性与论证性之间的往返变化，是索绪尔任意性原则向可论证性原则复归的第三步，由此完成了由任意性原则出发，在展开性证明的过程中不知不觉地向可论证性原则一步步复归。

## 四、任意性原则的展开反证了可论证性

语言符号的概念与音响形象之间的联系，索绪尔起初用惠特尼（W. D. Whitney）的“约定论”来解释，即认为只有约定才可以把不同的声音与不同的意义结合成一个统一体，但后来他认为，在结合成为符号之前，无论是声音还是意义两者都是非定形的，因此最终放弃了“约定论”而归结为“任意性”。荀子的“（王者）约定（民众）俗成”说与索绪尔的“任意性原则”并不等价。索绪尔的任意性原则要点归纳如下：

1. 任意性是说所指与能指之间没有自然的或内在的联系；

2. 相对任意性即可论证性；

3.绝对任意性即可论证性的转移或丧失;

4.在一种语言内部,论证性和任意性往返变化。

对所指和能指的关系不加历史性探讨,而以"任意性"一言蔽之,是任意性原则中论证的第一个失误。以不同语言系统之间能指和所指结合关系的差别来证明同一语言系统之内能指和所指结合关系的任意性,是任意性原则中论证的第二个失误。用共时的比例掩盖历时的溯源,在模仿的近似性与任意性之间划等号,把符号的历史演变性与不可论证性混为一谈,是任意性论证中的第三个失误。索绪尔把任意性一分为二,认为相对任意性即相对地可以论证,是任意性原则向可论证性原则复归的第一步,把绝对任意性解释为可论证性的转移或丧失,是任意性原则向论证性原则复归的第二步,把语言内部演化运动阐述为不断地由论证性过渡到任意性和由任意性过渡到论证性,则是任意性原则向可论证性原则的全面复归。根据《教程》中的"在任何时代,哪怕追溯到最古的时代,语言看来都是前一时代的遗产"(107页),"事实上任何社会,现在或过去,都只知道语言是从前代继承来的产物而照样加以接受,因此,语言起源的问题并不像人们一般认为的那样重要"(108页)这些前提,正如兹维金采夫所说:"不是可以得出另一个更为合理的结论来吗?既然语言永远作为前代的遗产出现在我们面前,而且在其当时存在的每一时期,语言'现有的这样'永远是由可论证的要素构成的体系,从而对语言的研究不能不以这种可论证性为基础。那么,关于语言符号的任意性的命题还有什么实践意义呢?"[1]

索绪尔提出的任意性与李葆嘉所主张的可论证性之异同,图示如下:

[1] B·A·兹维金采夫《普通语言学纲要》。

| | 索绪尔的任意性 | 李葆嘉的可论证性 |
|---|---|---|
| 含义（一） | 没有任何内在联系＝没有自然联系根基＝不可论证的（一编一章§2） | 有内在联系＝有自然联系根基＝可论证的 |
| 含义（二） | 相对任意性：<br>(1)句段关系中可切分；<br>(2)联想关系中可类比。<br>绝对任意性：<br>相对任意性的转移或丧失。<br>（二编六章§3） | 可论证性：<br>(1)同左；<br>(2)同左。<br>磨灭了的可论证性；<br>可论证性的丧失。 |

索绪尔说："我们对符号的任意性有一种非常敏锐的感觉，认为事情可能是这样。"（102页）这种"敏锐"的"可能"不是神秘主义，又是什么呢？因此，索绪尔的符号任意性原则实际上是个虚构性原则，除了给人们造成语言符号的形成过程不可捉摸的错觉，只有舍弃对语言符号的历时系统性研究之外，语言符号的任意性命题在语言学研究实践上没有任何价值。虽然共时语言学以语言的"静态"假设为前提，但是毫无必要割断活生生语言的动态真实。幸而，任意性原则的展开性阐述反证了语言符号的可论证性。

在人类语言的发展长河中，虽然语言符号具有自发任意性的倾向，但是语言符号的可论证性是其原生属性，而任意性只是其次生属性。人类所认识的世界就是语言中的世界，语言符号（符号个体音义）的可论证性原则和语法结构（符号配列关联）的临摹性原则，又包括在人类符号的"隐喻性"或"仿真性"这样一个更大的认知语言学研究领域之中。

## 第三节　语言符号的论证模式及其价值

在中国，引进并试图阐释索绪尔符号任意性原则的学者，一方面并未梳理索绪尔观点的来龙去脉，另一方面深受曲解了的荀子"约定俗成"说的以讹传讹的影响。在进一步澄清索绪尔任意性原则的同时，必须正面阐述语言符号的可论证性，探讨语言符号的论证模式及其价值。

### 一、任意性的含义及其对立性概念

读过《教程》的人对任意性原则的理解则不尽相同，或以为任意性"是指概念对于它的语言形式的多种选择可能性"[①]。索绪尔在《符号的任意性》中表述得很清楚，串联索绪尔自己的表述：任意性 = 没有任何内在的联系 = 没有任何自然联系 = 不可论证的。他提醒人们不要把任意性理解为"自由选择性"，"概念对于它的语言形式的多种选择可能性"不合乎索绪尔对任意性的解释。

索绪尔从任意性出发，到任意性一分为二；再从相对任意性解释为可论证性，到绝对任意性解释为可论证性的转移，一步步地走到了初始观点的反面。

或有人担心，否认了符号的任意性原则，"我们就无法解释世界诸语言之不同，无法解释语言符号的可变性与不变性"，"要是某语义必然由某语音来表示，何来各种不同语言，又何来比较

---

① 周静芳《我读〈国外对索绪尔符号任意性原则的批评〉》，《语言学通讯》（油印本），1985年第1期。

呢?”① 世界诸语言之不同,是在诸群体的历史性认知过程中逐步形成的。每种语言的主体编织着自己的语言网络,根据各自对事物特征的捕捉来确定事物的名称,语言的系统制约性又调整着事物的名称。在认知过程中,语主能动地、历史地创造着的语言符号,从理论上来说,都是可能论证的。或者是在能指与所指之间寻找自然的、内在的联系,或者是在这一能指与那一能指之间寻找联想关系中的可类比。一句话,可论证性可以通过追溯某个符号的最初构成来获得。换而言之,符号的可论证性是指在某一符号系统内(可能涉及与之有亲缘关系的系统)进行历史性溯源,所揭示出来的隐藏在语词符号内部的理据性。它只能存在于某一系统之中,所以索绪尔用不同语言系统之间的差别性来证明语言符号具有任意性,在方法论上是错误的。

一提到可论证性,或有人立即联想到“必然性”,认为可论证性就是在说“某语义必然由某语音来表示”,这可能纯粹是误解可论证性。索绪尔所提出的是“任意性”,而不是“必然性”的对立概念“偶然性”。根据索绪尔所提出的“任意性”,只能反推出“可论证性”或“理据性”——推不出“必然性”——而《教程》中所使用的对立性术语就是“可论证性”。一些教科书上将“任意性”的对立面限制为“必然性”,通过批评音义结合的“必然性”而证明“可论证性”的“荒唐”以避开了对“任意性”的证明,或者赞同“理据性”(词的内部理据即论证性)而批驳“可论证性”(可论证性即寻找词的内部理据),显然是驴头不对马嘴。

## 二、单纯性符号的两种论证模式

语言符号可分为单纯性符号和合成性符号,其论证方式也

① 周静芳《我读〈国外对索绪尔符号任意性原则的批评〉》。

可分为整体式论证与分解式论证。索绪尔在相对任意性中采用的句段切分与联想类比，就是针对合成性符号的分解式论证。然而，更值得探讨的是单纯性符号的整体式论证，因为这一研究更能明确语言符号的论证性。

根据能指和所指之间有无自然联系的标准，可以把单纯性符号分为"有自然联系"与"无自然联系"两类。对前者可采用直接论证模式，就某一具体能指和所指之间，用所指的特征来说明能指被确定的原因；对后者则采用间接论证模式，就能指和能指之间，用同一义类(特征性语义场)中的这一能指来说明那一能指被确定的原因。论证的基础就是这两个能指的各自所指，它们之间存在某种命名依据或词语理据上的共通之处。索绪尔在相对论证性中只提到合成结构的切分，至于单纯结构的派生孳乳却没有提到，然而，汉语中存在单纯词派生孳乳的大量现象，并且由此形成同源或同族字词类聚。那些自然联系似乎已经磨灭、暂时无从追索的符号，可以暂缓探究。

对于某一语言系统内符号的论证工作，可以从理论上设定为如下步骤：第一步，断代性截取，必须把某一系统的符号限制在某一时代以前。第二步，分离出合成词，对其进行分解式论证。第三步，在单纯词中分离出具有自然联系的一类，进行直接论证。第四步，根据"音近义通"的原则，多方验证，把相关的单纯性符号串联起来，组成一个个孳生系列或词族，以对之进行间接论证。依据间接论证模式，在能指与能指或名与名之间系源的研究[①]，大致有三种情况：

1.顺向系源，以本源词论证孳生词。如：梯、弟也。粉，分也，研米使分散也。

---

① 参见陆宗达、王宁《浅论传统字源学》，《中国语文》1984年第5期。

2. 横向系源，以孳生词论证孳生词。双项系源的如：胫，茎也，直而长似物茎也。多项系源的如：皠，鸟之白也；犨，白牛也；骓，马白额；鹤，色多白；翯，鸟白肥泽貌；颢，白首人也；皓，日之门也[1]。

3. 逆向系源，以孳生词论证本源词。如：皮，被也，被复体也。

只能采用相对的标准确定根词，即找不出本源词的源词就作为根词。根词的数量是有限的，有多少个孳生系列就有多少个根词。源于远古语言的根词，如果其中有些是象声词，可以从自然联系中去直接论证。先民以野果为食，野果落地，滚转有声，音近 * gulu，初民模仿以为名称(写作"果裸")。其后思维日精，语言益晰，从"果裸"这一语根中逐渐生分化出多种水果的具体名称。据研究，杏、梨、李、棠、桃、橘、柚…… 大多可能源于"果裸"[2]。

## 三、根词推断论证性的类比证明

那些一时还没有足够资料加以论证得名之由的，宁付缺如。从理论上，绝缘无佐证不等于不能论证，可论证性的磨灭并不表明这种可论证性原初从未存在过。对视觉能指——文字——的分析启示我们，听觉能指——语词——的可论证性比任意性接近事实真相。

首先，对文字的研究证明，这一系统内的可论证性符号转入另一系统内会成为不可论证性符号。英文字母是表音符号，字母"A"之形与语音[ei]之间没有内在的自然的联系，在这一系统

① 杨树达《积微居小学述林·释皠》，中华书局 1983 年。
② 任继昉《汉语语源学》，重庆出版社 1992 年

内,"A"是不可论证的无理性符号,似乎仅仅是约定。然而,如果追溯字母的由来,进入更早的符号系统,情况将有所不同。撇开中间过渡环节,直至古埃及图画文字系统中,即会发现这些字母脱胎的原始形态,与所指之间具有明确的象征关系。字母"A"象"牛头"之形,在古埃及是表示公牛或家畜的象形符号。公元前18世纪以后,闪米特人借用它做音符,后经腓尼基人传至希腊成为希腊字母,罗马人又将之改造为拉丁字母方逐步流布欧洲。经过几千年的多次转移借用和形体演变,使用这种字母的人群再也不必了解原初的象征性了,仅仅把字母看作纯粹的音符。如果不是考古学的发掘和古文字的研究,面对现代使用的这些字母符号,自然只会认为它们具有不可论证性。基于这一系统内的有理性符号转入另一系统内,将成为无理性符号的这一事实,有把握推论,所有根词原初都是可论证的。只是听觉能指不能像视觉能指那样具有超时空性,它的有理性在历史发展的长河中易于泯灭罢了。

其次,在汉字系统内,古今文字的演变表明符号的发展呈现由象征性向非象征性过渡的趋势。以"刀"为例,图示说明如下:

| 序号 | 字　例 | 造字法 | 能指和所指关系 | 备　注 |
|---|---|---|---|---|
| 1 | 刀,兵也。 | 象形 | 能指和所指之间存在自然联系 | |
| 2 | 刃,刃坚也,象刀有刃之形。 | 指事 | 能指和所指之间存在有自然联系 | 在刀的符号上加点。 |
| 3 | 初,始也。从刀衣,裁衣之始也。 | 会意 | 由两个有自然联系的能指合成。 | 以刀裁衣。 |

续表

| 序号 | 字 例 | 造字法 | 能指和所指关系 | 备 注 |
| --- | --- | --- | --- | --- |
| 4 | 刀，小船。《诗·河广》："曾不容刀" | 同源假借 | 借刀为舠，字表标音，词语同源。 | 据语源，小船弯形似刀，故名之。 |
| 5 | 刀，古代钱币的一种。 | 同源假借 | 借刀做名词，字表标音，词源相同。 | 据语源，钱币形似刀，故名之。 |
| 6 | 刀，计算纸张的单位 | 同源假借 | 借刀做量词，字表标音，词源相同。 | 据语源，一沓纸用刀切齐，故名之。 |
| 7 | 苆，苇花也。从草，刀声。 | 同源形声 | 以刀做声符，兼表语义。 | 据语源，苇花似刀之形，故名之。 |
| 8 | 魛，鱼的一种。 | 同源形声 | 借刀做名词，字表标音，词源相同。 | 据语源，体长而侧扁似刀，故名之。也作刀鱼。 |
| 9 | 削（鞘），鞞。从刀，肖声。 | 形声 | 以刀做义符，有自然联系。 | 表明此为装刀之器。 |
| 10 | 叨，唠叨。 | 假借 | 以刀做声符，唠叨中的"叨"为音节字。 | |

汉字的这种发展趋势不仅从造字或用字方式的演进中可以看出，而且从文字形体的演变中也可以看出。从最原始的图画文字，经历甲骨文、金文、篆文，直至隶书，形体越古象征性越强。楷书已经变成非象征化的"横竖撇捺"了，太阳之"日"已经变形为窗户格子了。如果没有《说文》的流传和古文字的发现，仅仅依据楷书形体，"汉字是任意的、无理性的"观点盖有人提出。不

论是这一系统内的有理性符号转入另一系统成为无理性符号,还是表意文字系统内的符号发展呈现从象征向非象征过渡的态势,都表明视觉符号系统中的基本符号是可以追溯论证的。据此类比推断,听觉能指中的根词也具有可追溯的论证性。

根词的论证性问题牵涉到现代语言的起源问题。研究表明,直至约四万年前,人类才具有越来越完备的抽象思维能力,以有声语言为其思维形式。有什么样的神经系统就有什么样的思维,这种原始语言必然受制于当时神经系统的认知能力。思维的基本材料是通过感觉器官把外界刺激转换成的、神经系统中的某种感觉形态或表象系统,而认知的基本规律就是保持认知的材料、过程、结果与物质存在必须趋同的组合同构律。在这一规律的支配下,听觉表象总是力图使语音和语义尽可能形成某种同构,因此"象声"不仅是人类认识事物的基本途径,而且是语言产生的必由之路,这是人类在漫长发展中的唯一选择。与此同理,文字必然起源于"象形"。象声和象形是人类所具有的符号化认知能力的两大渊源。在人类的认知活动中或形成语言符号的过程中,存在一些不可摆脱的历史条件:历史主体的认识能力、具体客体的显性特征、具体的认知背景、已经存在的符号系统。

唐代无名氏《无能子·纪见第八》说:"且万物之名,亦岂有自然著者? ……皆妄作者强名之也。人久习之不见其强名之初,故沿之而不敢移焉。"[1] 在无名氏的"强名论"中,命名的主观随意性得到充分的发挥,与索绪尔的符号任意性原则似曾相识。归根结底,符号任意性原则的要害就在于将语言主体在具体的认知活动中的历史主动性一笔抹杀,用"共时"之剑斩断了语言

① 转引自齐佩瑢《训诂学概论》。

的历时源流关系。

## 四、语言符号论证性原则的价值

与符号任意性原则的相反，符号的可论证性原则认为：通过历时地追溯名与实或名与名之间的关系来探索语言符号的论证性，对复合结构采用分解式论证模式，对单纯结构采用整体式论证模式，那些明显的有自然联系的符号具有直接论证性，那些隐晦的（置于词族之中可以显现的）有内在联系的符号具有间接论证性，无可追溯的、论证性磨灭了的根词则具有推断论证性。

语言符号的可论证性原则具有重要的理论价值。可论证性为词源学、字源学、词汇系统的研究，为词的内部形式或理据性探讨提供了理论基础。不必既认定"同源的派生词的音义，由于都是从根词早已结合在一起的音义直接或间接发展而来的，因此带有历史的、可追溯的必然性"，又交代"'约定俗成'说明了词的音义联系不是天然的，有机的，而是偶然的"，更不必推定"从词汇的总体说，约定俗成仍是音义联系的总规律，音近义通不过是词汇发展某一方式所体现的个别规则，二者在理论上是不可列入同一层次看待的。"[1] 在单纯词为主导的古汉语符号系统中，由推衍命名方式孳生语词是普遍现象，音近义通非个别规则。音义总规律也许可以这样表述：某一语言系统内，能指和所指之间的关系是历史地形成的，二者之间的两种联系——象征性与非象征性——都是可以论证的，用历时观点考察，中间可能存在一个模糊过渡段。

可论证性抹去了命名方式或音义结合关系上的神秘主义色彩。初民们在自己的历史活动和认知过程中，绝不会对自己的

[1] 陆宗达、王宁《浅论传统字源学》

认知对象的命名漠不关心,他们具有认识的能动性,符号具有内在的系统性。命名就是认知和抽象,名称牢牢烙下了认知的痕迹。当代人给新事物的命名过程可以看作是初民命名的模拟实验,新词的产生总是不断地向任意性原则提出挑战。可论证性理论能够提高学人以科学的态度处理传统词源或字源学研究资料的勇气。可论证性理论导同源词的研究,进而从事根词的特征、形成条件诸方面的探索。作为最古老的语言细胞的根词,与语言中较为稳定的因素如构词和构形法存在着密切的联系。根词与古音构拟也具有特殊关系,在语言亲属关系的确定和共同原始语的构拟中,不同语种之间的根词的比较研究比基本词汇的研究更有价值。

依据语言符号的可论证性原则,在词族研究、根词研究的基础上,结合其他领域内的成果,可以进而探讨初民的心理机制、认知模式、思维方式及命名规律。初民命名规律的系统分析研究可形成新的学科——原始命名学。反之,词源学、字源学、命名学、历史比较语言学等方面的研究,又会进一步证明语言符号可论证性原则的真理性。

# 第四章　静态语言学的来源和心理索绪尔

## 第一节　静态语言学的三个直接来源

### 一、《教程》索绪尔和学术索绪尔

1916 年，索绪尔《普通语言学教程》① 的出版，被喻为语言学领域的"哥白尼革命"。回首百年来的索绪尔和《教程》研究，有几个基本问题长久以来却没有解决：1. 索绪尔静态语言学的来源是什么？比如叶尔姆斯列夫（L. Hjelmslev）就曾说过："很难说，索绪尔的观点是如何在思想中具体形成的，而我个人的理论和方法，许多年以前在我接触索绪尔的观点之前就已经形成了。"② 2. 是什么内在动力驱使索绪尔去建构静态语言学体系？3. 索绪尔静态语言学的价值到底是什么？显然，这些问题就《教程》而研究《教程》是不可能解决的，只有将索绪尔理论置于宏观学术背景上并借助其他资料深入到索绪尔的内心世界，才有希

---

① 《普通语言学教程》（高名凯译），商务印书馆 1982 年。以下简称《教程》，引文所注页码即中译本页码。

② L. Hjelmslev: *Structural Analysis of Language*，转引自冯志伟《现代语言学流派》，73 页，陕西人民出版社 1987 年。

望取得逐步进展。

被称之为"结构主义"鼻祖的索绪尔，其理论是通过《教程》而得以传播的。通常所认识的"现代语言学"理论，就是《教程》中所阐述的理论。作为《教程》索绪尔：

1. 区分语言和言语，在语言语言学和言语语言学之间选择前者；

2. 区分语言学的内部和外部，在内部语言学和外部语言学之间选择前者；

3. 区分内部语言学的静态和动态，在静态语言学和动态语言学之间选择前者；

4. 区分语言学的形式和实体，在形式和实体之间选择前者。因此索绪尔的"现代语言学"可以定义为：以语言符号系统为对象的、内部的、静态的、形式语言学。正如《教程》结尾的一句名言所揭示的那样："语言学的惟一的、真正的对象是就语言和为语言而研究的语言"。因为时间总是变动的，因此用"现代语言学"作为这一流派的名称必然含混不清，而"结构主义"不是索绪尔本人提出来的①，虽然这一来源于心理学格式塔理论的"结构"，作为一种被分离的整体，与索绪尔的"语言是一种表达观念的符号系统"的"系统"或"语言是形式而不是实体"中的"形式"含义相当。合适的术语可能是《教程》中的术语：与言语的语言学相区别的"语言的语言学"（40 页）、与演化语言学相区别的

① ［苏俄］康德拉绍夫（H. A Кондрашов）《语言学说史》（杨余森译，武汉大学出版社 1985 年）提出，结构主义这一术语是荷兰语言学家波斯（H. Pos）在 1939 年第一次使用的（145 页）。今考波斯在 1939 发表《结构主义的观点》。又据钱军《结构功能语言学》（134 页），雅柯布森（Roman Jakobson）在《浪漫的泛斯拉夫主义》（1929）中指出："如果要囊括当前各种科学的主导思想，再没有比结构主义更贴切的术语了。"最先创用"结构主义"这一术语。

“静态语言学”(117页)或与历时语言学相区别的“共时语言学”(143页)。如果进一步思考，那么“语言的语言学”是以研究对象作为限定，并非代表索绪尔的研究方法，而语言的共时态，在索绪尔的概念中其实就是“现时态”或“当时态”。因此，虽然语言的“静态”在语言的长河中并不存在，但是可以作为科学思维的假设代表索绪尔语言学中的核心观点，因此以“静态语言学”最为适宜，索绪尔本人也习惯于这么称呼。

20世纪50年代起，经过瑞士戈德尔(R. Godel)、意大利毛罗(T. de Mauro)、瑞士恩格勒(R. Engler)和俄国斯留萨列娃(Н. А. Слюсарева)对新发现的索绪尔的手稿札记、学生课堂笔记、回忆录和信件等材料的研究，最终导致了毛罗所说的“索绪尔的语言学思想的重新发现”，即出现了一个与“《教程》索绪尔”不完全相同的手稿索绪尔[①]。与索绪尔的手稿和学生课堂笔记的对照，不难看出沙·巴利等编撰者在整理《教程》过程中对课堂笔记及有关讲稿的增删改拼痕迹。甚至有人认为，《教程》结尾的那句“灾难性”名言，也是编者擅自所加。然而，既然我们无法否认这句名言确实是对《教程》内容最精辟的总结，也就不能排除索绪尔生前曾经对他的学生确实说过这样类似的话。不见诸《札记》的话，只表明残存的《札记》中没有记录，并不能作为索绪尔生前没有说过的证据，这一点是容易理解的。

在发现者看来，“手稿索绪尔”的特点是：1. 索绪尔主张既研究语言，也研究言语。索绪尔第三次讲授普通语言学的提纲手稿中，计划把课程分为三部分：第一部分标作《种种语言》，第二部分标作《语言》，第三部分他许诺讲授“个人的言语活动能力

---

① 戚雨村《索绪尔研究的新发现》，《现代语言学的特点和发展趋势》，外语教学和研究出版社1997年。

及其表现"。因此有人认为,言语语言学的不能很快建立不能归咎于他。2. 索绪尔强调共时分析的重要性,但并不否定和排斥历时分析。3. 索绪尔说过:"语言学的内部研究和外部研究,……语言学这一术语首先包含着二者统一的思想"[①],由此可见他并没有否定外部语言学。

然而,必须看到:1. 虽然"手稿索绪尔"并不反对且计划讲授言语的语言学,但是不可否认,言语语言学的不能很快建立确与"《教程》索绪尔"密切相关。2. 虽然早期索绪尔是一个富有创建的历史语言学家,但是在《札记》中,同样写下了索绪尔对历时研究深为不满的议论:"语言学家要研究的是语言态[②],他不需要理会导致目前语言态的历史事实,他应该把历时研究置于不顾。……历史的干预只能歪曲他的判断",[③] 又大声疾呼:"我们必须做出反应,抵制老学派的邪道……"[④] 3. 虽然索绪尔在《教程》和《札记》中都没有否定外部语言学,但也没有从事外部语言学研究。

由此可见,索绪尔的新发现者忽视了一个显而易见的事实:问题的关键不在于索绪尔本人有没有否定言语的语言学、历时语言学和外部语言学,而在于不排除这些"杂七拉八"的干扰,《教程》的静态语言学体系就无从建立。[⑤]不容否认,正是以《教

---

① 恩格勒《评注本》,59 页。转引自《索绪尔研究的新发现》。

② 指语言静态或共时态,即研究者面临的"当时态"。

③ 毛罗《评注本》,117 页。转引自许国璋《许国璋论语言》,151 页,外语教学和研究出版社 1991 年。

④ 戈德尔《稿本溯源》,252 页。转引自许国璋《许国璋论语言》,105 页。

⑤ 由于不了解这一关键,以致于、美]R·H·罗宾斯《简明语言学史》(许德宝等译,223 页,中国社会科学出版社 1997 年)写道:"奇怪的是,在共时语言学中极重视结构主义思想的索绪尔,竟然否认结构跟历史研究的关系。"

程》为形式的索绪尔，才使其精辟论点引起了欧美语言学界的广泛瞩目，也正是"《教程》索绪尔"在对后几代语言学家产生影响的同时限制了语言学研究的宽阔视界，才导致了言语语言学、外部语言学顶着结构主义的压力而艰难诞生的状况。索绪尔的重新发现不过表明，批评、修正和企图突破静态语言学的那些"邪说"在"手稿索绪尔"中也存在过。人们在对《教程》编者表示不满的同时，切勿忘记关键的一点：如果没有《教程》，索绪尔不会被奉为结构主义语言学的鼻祖；也正是有了《教程》，索绪尔因此才有可能被《教程》阅读者来"误读"。

除了这两个"索绪尔"，还有尚未展开充分研究的学术索绪尔。静态语言学的讲授体系即《教程》索绪尔，静态语言学的理论来源即学术索绪尔。通过揭示"学术索绪尔"或索绪尔语言理论直接渊源的寻求，可以了解索绪尔为何成其为"《教程》"和"手稿"的索绪尔及其在理论语言学史上的切实地位。

关于索绪尔的生平和学术风格，巴利的《序》和卡勒(J. Culler)的《索绪尔》[1] 做了一些叙述。关于索绪尔语言论的来源，学者们做了一些探讨。《教程》采用了洪堡特(W. von Humboldt)反对唯理语法所提出的论点(125 页)，或指出索绪尔的理论与洪堡特的"价值"或"系统"论述有渊源关系，惜未免远绍。或依据《教程》中叙述，认为索绪尔的语言价值说受以华尔拉斯(Warlars)等人为代表的瑞士正统经济学派理论的影响比较深，从中吸取了"一切研究价值的科学的内在二重性"的观点(117 页校注)，但二分法仅是取舍方法。关于索绪尔静态语言学理论的直接来源，虽有只言片语触及却始终未见深入而全面的阐明。然而，有可靠证据表明，《教程》的理论体系与法国社会学家德克

[1] [美]卡勒《索绪尔》(张景智译)，中国社会科学出版社 1989 年。

海姆(E. Durkheim,1858—1917)的社会学理论,与波兰—俄国语言学家博杜恩·德·库尔特内(Baudouin de Courtenay,1845—1929)及其弟子克鲁舍夫斯基(H. B. Крущевский,1851—1887)的语言学理论,还有美国语言学家惠特尼(1827—1894)的语言符号观存在直接的渊源关系。

## 二、德克海姆的社会理论:索绪尔语言论的哲学基础

卡勒在英文本《教程·引言》(1974)中称弗洛伊德(S. Freud)和索绪尔、德克海姆为近代三大思想家,其生年各为1856、1857和1858年,分别创立了现代心理学、现代语言学和现代社会学。俄国语言学家兹维金采夫和布达哥夫认为,索绪尔语言论的哲学基础是德克海姆的社会学。索绪尔的学生梅耶(A. Meillet)等提出,德克海姆的《社会学研究准则》出版在1895年,而索绪尔1891年已经去了日内瓦,不大可能受其影响。然而,波兰学者罗雪夫斯基在法国心理学杂志发表了《论社会学与语言学的关系:德克海姆与索绪尔》(1933),证明索绪尔对德克海姆与塔尔德(G. Tarde)之间的社会学论争很感兴趣,在形成自己的语言理论中有所反映。

1959年,方光焘在《涂尔干的社会学与索绪尔的语言学理论》[①]一文中认为,即使1891年[②]以后,索绪尔已不在巴黎,同样可以经由学术交流了解德克海姆的理论。根据方光焘的研究,再结合其他材料比照,索绪尔语言论的哲学基础是德克海姆的社会学理论确凿无疑。

---

① "涂尔干"是Durkheim的另一汉译;又许国璋译为"杜尔淦",张景智译为"杜尔克姆"。

② 方光焘《语言学论文集》中误为"1881年",商务印书馆1997年。

（一）对言语活动与对社会事实抽象的一致性

德克海姆认为："社会事实是在个人之外、对于个人有一种强制力的行动、思维和感觉的样式"；社会事实是个人的集合；社会事实通过个人表现，这种个人表现依靠个人生理、心理的构造。索绪尔认为"语言"本质上是社会的，独立于个人之外的；语言系统是个人言语的集合（1+1+1+……=1）；"共时语言学研究社会集体意识所感到的心理的逻辑的关系"。

（二）对语言和对心理社会属性认同的一致性

在心理学论争中，塔尔德强调个人意识，德克海姆强调社会意识。索绪尔认为，语言是社会的，言语是个人的。他以语言的社会性为基础建立语言的语言学，但又对言语的语言学适当让步。

（三）对社会规律两个基本特征认定的相承性

《教程》说："任何社会规律都有两个基本特征：它是命令性的，又是一般性的"，中译本校注：索绪尔对于社会规律的两个基本特征是按照法国社会学家涂尔干的理论来理解的。（132 页）涂尔干认为，一切社会规律必须符合两个基本原则：强制性和普遍性。索绪尔所说的命令性等于强制性，一般性等于普遍性。

（四）对社会和对语言静态与动态划分的相承性

德克海姆继承了孔德（A. Comte）的社会静力学和社会动力学的区分，把社会学分为静态社会学和动态社会学。与之相应，索绪尔把语言学分为静态语言学与动态语言学。

（五）对语言学属于社会心理学看法的相同性

德克海姆认为语言学是社会心理学的一部分，索绪尔认为符号学（语言学是其典型）是社会心理学的一部分。德克海姆认为社会学与心理学不可分割；索绪尔认为语言学与心理学不可分割，并提出语言符号不是事物与名称的结合，而是音响形象与

概念的心理结合。

尽管结构主义的追随者们摆脱了社会心理学说,但是索绪尔认为自己的理论与社会心理学密切相关。苏俄语言学家大都把索绪尔归人"社会心理学派"或"社会语言学"流派。必须指出,这一以心理学说为哲学基础的"心理学派"与当代用心理学的方法研究语言的"心理语言学"不同;这一以社会学说为哲学基础的"社会学派"与当代用社会学方法研究语言的"社会语言学"不同。依据《教程》的观点,心理语言学和社会语言学属于外部语言学。令人感到有趣的是:以社会心理学为哲学基础的静态语言学系统建构以后,却导致了语言学的研究摈弃心理学与社会学具体方法的后果。

## 三、博杜恩的语言理论:索绪尔语言论的理论框架

作为创建了喀山学派的博杜恩,在索绪尔以前很早就提出了"现代语言学"的一系列原则和理论体系。博杜恩 1870 年获得彼得堡大学比较语言学硕士学位和莱比锡大学哲学博士学位;1874 年获得彼得堡大学比较语言学博士学位(1873—1875,索绪尔在日内瓦公立高中上学)。1887 年当选为克拉科夫科学院院士,1897 年当选为彼得堡科学院通讯院士,与叶斯泊森和梅耶等都有通讯联系。

作为一位富有创造天才的探索者的博杜恩,正如他的学生谢尔巴(Л. В. Щерба)所说:"他不是任何人的学生,也不属于任何学派,他自称是自修者"。博杜恩一生中写过 600 多部篇论著,是一位勤奋而多产的著作家。主要著作有《对语言学和语言的几点总的看法》、《语音交替理论探索》、《语言学札记和箴言》和《语言学引论》。他的普通语言学、比较语言学研究,不仅对当时的历史语言学、而且对后来的结构主义有着重要的影响。他

的有关论述受到20世纪中叶兴起的心理语言学、社会语言学、对比语言学和应用语言学界的重视。由于他所处的时代还是热衷于历史语言学的时代,晚年又没有出版全面阐述自己理论的著作,他的声名一直为《教程》索绪尔所掩没。直到20世纪下半叶,国际语言学界才逐渐认识到他的独创性观点和富有洞察力的见解。

作为自然主义语言学派代表人物施莱歇尔(A. Schleicher)的学生,博杜恩不赞同把语言归结为自然科学。他认为:"这两种理论(谱系树说和波浪说)都经不起推敲。因为一方面,它们都是从语言脱离人而存在这一前提出发的,另一方面,它们都没有充分考虑到语言现象的复杂性"。博杜恩的学术思想以德国哲学家赫尔巴特(J. F. Herbart)的个人心理联想学说(心理活动是基本感觉单位的表现)为基础,同时又吸收了拉撒鲁(M. Lazarus)的民族心理学观点(心理研究应着眼于社会整体),从社会学的角度来解释心理主义,认为"人类语言的本质完全是心理的。语言的存在和发展受纯粹心理规则的制约,人类言语或语言中的任何现象,同时又是心理现象"。他所创立的语言论具有浓厚的社会心理学倾向,始终贯穿着一条线索:就是通过心理机制的分析,对语言规则、语言功能和语言发展做出解释。他主张把心理学和社会学糅和在一起作为语言学的基础,并提出语言学属于"心理一社会科学"的看法。

早在1870年(索绪尔13岁),在《教程》出版(1916)前近半个世纪,博杜恩已经提出语言和言语的区分。与索绪尔后来的区分不完全相同,他区分为抽象的民族语言、具体表现为个人语言的民族语言和在个人语言意识中反映的中间语。1876年(索绪尔进入莱比锡大学),他提出语言的静态和动态的区分(索绪尔1897年笔记中出现历时态、共时态和特殊共时态的术语)。

他认为语言现象之间的联系,不仅存在于历时的发展之中而且存在于共时的状态之中,但不主张共时分析应优先于历时分析。在他看来"语言中没有静止不变的东西",静态与动态并不截然分开,静态从属于动态,这种"静态寓于动态"的论断比《教程》索绪尔切合语言事实。1876—1877年(1876年5月索绪尔加入巴黎语言学会,10月进入莱比锡大学),博杜恩提出语言是一个系统的学说,语言是其组成部分(语音、语义和形态)的总和。他的学生克鲁舍夫斯基发展了他的思想。在1883年的博士论文《语言学概要》(1881年起,索绪尔在巴黎高等学院任教)中,除了提出"词是事物的符号"、"语言……是一种符号系统",不但强调符号的任意性问题,而且主张这个系统既可以"在同时共存中"分析,又可以"在连贯序列中"分析。同时,克鲁舍夫斯基还提出语言符号的类比联想和邻接联想规则,比索绪尔(1906年第一次讲授普通语言学)的联想关系和句段关系早20多年。作为音位学的创始人,博杜恩(1881)提出必须明确区分音素和音位,但是他不止一次地指出,这两者的区分是克鲁舍夫斯基倡议的。受博杜恩的影响,布拉格学派特别注重音位研究,并且公开反对索绪尔对共时和历时的绝对割裂,集中体现在流亡布拉格的俄国语言学家特鲁别茨柯依(Н.С.Трубецкой)的《音位学原理》中。《教程》对音位学并无实质性建树,与其说布拉格学派接受了索绪尔的理论,不如说发展的是喀山学派的观点[①]。

1881年到1882年,作为巴黎语言学会永久会员的博杜恩,与时任学会秘书的索绪尔几次在会议上晤面。博杜恩不仅把他

① 赵世开主编的《国外语言学概述——流派和代表人物》(334页,北京语言学院出版社1990年)说:特鲁别茨柯依成功地把索绪尔的对立论和博杜恩的功能观运用到语音学领域,成为布拉格学派音位学理论的主要代表。

和克鲁舍夫斯基的论文提交给学会，而且索绪尔与博杜恩之间书信往来密切，博杜恩曾把他的和克鲁舍夫斯基的论著寄给索绪尔。索绪尔高度评价过喀山学派的思想，在一篇札记中写道："博杜恩和克鲁舍夫斯基比其他任何人更逼近于从理论上理解语言的意义，他们没有溢出纯粹语言学的范围"。

由于博杜恩的语言学理论的早已存在，《教程》传到苏俄时，语言学界认为并无多少新意。无怪乎他的学生谢尔巴不得不说："1923 年，当我们在列宁格勒收到索绪尔《教程》一书时，使我们感到惊讶的是，索绪尔与我们所熟悉的原理在许多地方惊人的相似。"当时的博杜恩已经回波兰华沙，作为爱国者和政治活动家的他(1922 年被推选为总统候选人)已无暇顾及语言学研究。毫无疑问，博杜恩是《教程》索绪尔语言理论的先驱，他和克鲁舍夫斯基才是"现代语言学"，不仅指静态语言学，包括应用语言学和比较语言学等的真正开创者。

## 四、惠特尼的符号学说：索绪尔语言论的符号原则

索绪尔称赞博杜恩、克鲁舍夫斯基和惠特尼是他最推崇的同时代的三位语言学家，1894 年索绪尔还专门起草过一篇纪念惠特尼的文章。雅柯布森在《世界语言学界对惠特尼语言科学原理的反应》(载 M. Siliverstwin 主编《惠特尼论语言》，1971)一文中，论述过索绪尔的语言理论与惠特尼学说的关系。

惠特尼认为，语言学属于人文科学，信息交换功能是语言的基本功能。他写道："语言……是说出来、听得见的符号"，由此在语言的"符号性"的基础上确立了符号的"约定性"与"任意性"、"不变性"与"可变性"等性质。《教程》中提到惠特尼对欧洲语言学的影响，并且多次引用惠特尼的符号学观点以阐述《教程》中的符号任意性原则：

发出第一次冲击的是《语言的生命和发展》(1875)的作者美国人惠特尼。 (24 页)

惠特尼就把语言看作一种社会制度……在主要论点上,我们觉得这位美国语言学家是对的:语言是一种约定的东西,人们同意使用什么符号,与这符号的性质是无关紧要的。 (31 页)

为了使人感到语言是一种纯粹的制度,惠特尼曾很正确地强调符号有任意的性质,从而把语言学置于它的真正的轴线上。但是他没有贯彻到底,没有看到这种任意的性质把语言同其他一切制度从根本上分开。 (113 页)

符号的不变性和可变性。 (107 页)

据徐志民论述,索绪尔起初是用惠特尼的"约定论"来解释语言符号,即认为只有约定可以把不同的声音与不同的意义结合成一个统一体;但他后来认为,无论是声音还是意义,在结合成符号之前,两者都是捉摸不定的。因此,索绪尔最终放弃了"约定论",而归结为"任意性",但并没有完全放弃约定论[①]。符号的"约定性"和"任意性"是对立的统一。惠特尼是以"约定论"为主导的,《教程》索绪尔对约定性和任意性都是肯定的,但是他批评惠特尼,没有把这种可以将语言同其他一切制度从根本上分开的任意性贯彻到底。在索绪尔符号学原理中,选取了"任意性"作为语言符号的基本原则,而后又通过把任意性分为绝对和相对,从而以"相对任意性"这一术语又把"约定性"纳入。符号的"不变性"和"可变性"也是对立的统一。同样,在索绪尔的符号

① 徐志民《欧美语言学简史》,学林出版社 1990 年。

学原理中，"符号的任意性本身实际上使语言避开一切旨在使它发生变化的尝试"。(109 页)在索绪尔看来，符号的任意性决定了符号的不变性，并且指出了阻碍这种变化的 4 种因素。

索绪尔的语言符号学是在惠特尼的语言符号观基础之上改造建构的。索绪尔断言："符号的任意性原则没有人反对，但是发现真理往往比为这真理派定一个适当的地位来得容易"(103 页)，惠特尼发现了"任意性"但没有"派定"这一真理适当的位置，而后者比前者困难。把任意性确定为符号的基本原则，索绪尔对此颇为自鸣得意。

以法国学者德克海姆的社会心理学理论为哲学基础，以波兰—俄国学者博杜恩和克鲁舍夫斯基的语言学理论为理论框架，以美国学者惠特尼的语言符号观为基本原则，排除语言的动态研究，把静态研究放在惟一的位置，这就构成了索绪尔作为讲稿、而他的学生视为遗著出版的《普通语言学教程》这本著作中的"静态语言学"体系。

## 五、当代世界语言学对结构主义的全面超越

结构主义最初以日本为媒介传入中国，小林英夫将《普通语言学教程》(1928)翻译成日文，书名为《言语学原理》，王古鲁《言语学通论》(1930)列为参考文献，"索绪尔"译为"苏秀尔"。1931 年，在法国里昂大学归来的方光焘，首先讲授索绪尔和房德里耶斯的理论。"文法革新讨论"中，陈望道介绍了一些基本术语，方光焘提出了以"结合关系"为基础的"广义形态说"。

既然讨论中已经引介，为什么索绪尔理论在 30 年代的中国反响不大？当时几乎没有一篇运用索绪尔理论方法研究汉语的论文。甚至可以说，即使在 80 年代中译本出版以后，读完《教程》的人也不多。面对这样一本内容庞杂(语言学史、语音学、符

号学、共时语言学、历时语言学、地理语言学、史前史语言研究、语言类型学)、理论抽象(社会心理学、关系思维、价值系统)。由讲义整理而成的书,读懂、读通并非易事。40年代的中国正处于热中"三品说"的时代,而《教程》明显缺乏研究的操作性。作为结构主义的索绪尔,是借助布拉格学派的音位分析法和美国描写学派的语法分析法的广泛流传而名声大振的。

尽管出版了以布龙菲尔德(L. Bloomfield)《语言研究导论》为理论的张世禄《语言学原理》(1931),陆志韦(1938)把描写主义运用于汉语单音词词汇研究,王力的《中国语法理论》(1945)吸收了"向心"和"背心"术语,然而全面运用描写主义分析汉语事实的却是赵元任的 *Mandarin Primer* (1948)。50年代初,中国语法学界通过赵元任的研究为媒介而接受了来自美国的结构主义,*Mandarin Primer* 所体现的描写主义在中国语法学界占据主导地位长达30年。

描写主义在中国的遭遇是曲折的。在全盘苏化的背景之下,西方语言理论被斥之为资产阶级加以"批判"。随着历史性的解冻,苏联重视对结构主义的研究,中国语言学界也开始介绍。中苏两党争论的公开化,使这些介绍还没有充分展开就又受到批判。虽然如此,中国语言学界在批判西方的同时,只要可能就偷运描写语言学方法。《语言学资料》在主要介绍苏联语言学的同时,也译载了相当数量的美国描写主义代表作,布龙菲尔德的《语言科学的一套基本原理》、弗里斯(C. C. Fries)的《布龙菲尔德学派》等翻译,都促进了对描写主义的了解。50年代末到60年代初,出现了与之有关的"语言和言语问题的讨论"和"关于《说"的"》的讨论"。60年代,翻译出版了萨丕尔(E. Sapir)的《语言论》和弗里斯的《英语结构》,桂灿昆发表了《索绪尔语言学理论简述》、《索绪尔思想对美国结构语言学派的影响》

等文。然而，1965年的《中国语文》对结构主义进行大批判。虽然朱德熙运用了描写主义方法，但在《关于〈说"的"〉》中却不得不否认其方法是描写语言学的……

结构主义把语言看作一个封闭的静态系统，因此对语言各个领域的影响及其结果各不相同。音位系统是封闭的，布拉格学派建立的音位学理论和操作方法与中国传统等韵学的精神一致，因此语音系统（包括方言音系和历史音系）的成就最大。词汇系统是开放的而研究难度最大，《教程》以"符号的任意性原则"避开了其研究，汉语词汇研究依然是一盘散沙。因为语法系统相对封闭，美国描写学派发明了分布理论和层次分析法，加上许多人误以为"只有研究语法才是语言学家"，故趋之若鹜且成果明显。都是，由于结构主义区分语言和言语，割裂共时和历时，注重关系而舍弃实体，强调形式而排斥意义，重视描写而轻视解释，由于赖以建立西方语言理论的印欧语事实与汉语事实的参差龃龉，在促进汉语研究的同时造成了一系列负面影响。1996年，王士元先生应邀来访，我曾请教：美国语言学界讨论不讨论索绪尔？王先生含笑告之不讨论。早在1963年就译成中文、1980年才得以出版的《普通语言学教程》却为大陆语言学界掀起了一股索绪尔热。在反思和重建的潮流中，出现了以《语言文字应用》为主要论坛的"90年代语言观讨论"。与世界学术潮流相一致，90年代的中国语法学界终于实现了从一统独踞到多元并存的历史性转折，体现了当代世界潮流所倡导的"学术自由"的宽容性和"思想多元"的价值观。

虽然从《教程》出版以后，以布拉格学派、哥本哈根学派与美国描写主义学派为标志，结构主义成为欧美语言学的主流，但是人们也在以多种方式修正《教程》的一些观点，批评首先来自于内部。早在1928年海牙召开的第一次国际语音学会议上，布拉

格学派就发表声明，公开表示不同意索绪尔共时与历时绝对区分的观点，同时主张结构—功能语言观。即使在结构主义最盛行的时期，欧洲也从未成为结构主义的一统天下。除了叶斯泊森的新传统语法，英法功能学派保持独立思考，不愿为了维护一种既定理论而抛弃语言事实。马尔丁内（A. Martinet）一针见血地揭出静态语言学的要害："科学研究首先的要求，就是不能因为方法上的苛求而牺牲研究对象的完整性。"但是由于《教程》索绪尔影响太大，要克服人们头脑中的思维定势实在举步维艰，正如马尔丁内所说："功能语言学获得的任何进展，不论在过去还是将来一个时期内，都是逆着潮流的。"[①] 从50年代开始，以乔姆斯基（N. Chomsky）革命为导火线，当代语言学理论才突破索绪尔《教程》的彻底束缚，最终导致了静态语言学的全面超越。

（一）语言和言语区分被二者并容沟通所取代

伦敦语言学派的创始人弗斯（J.R. Firth，1890—1960）吸取了马林诺夫斯基（B. Malinowski）"意义是情境中的功能"的原则，不同意语言和言语的区分。弗斯区分了"结构主义语言学"和"结构语言学"而倾向性于后者。弗斯认为，语言是一个"社会过程"，是"人类生活的一种方式，并非仅仅是一套约定俗成的符号和记号"[②]，语言的异质性和非联系性，要比大多数人愿意承认的还要严重得多。新弗斯学派代表人物韩礼德（M. A. K. Halliday），接受了叶尔姆斯列夫的语言和言语二者并容的观点。他认为语言是"社会行为"，主张借助社会学理论来观察语言的使用，最终创立了"系统功能语法"。

① 《研究语言本身的语言学》，《语言学译丛》(1)，中国社会科学出版社。

② 转引自冯志伟《现代语言学流派》，159页。

（五）“线条性”只是语言结构的组合表象而“层次性”才是语言结构的构造顺序

语言的线条性只是语言结构的组合表象，美国描写语言学派发现语言结构的构造层次性，提出直接成分分析法。

（六）“组合关系”和“聚合关系”被“结构关联”所包容

组合和聚合关系建立在静态语言学的线性结构理论之上，“结构关联”是动态语言学的理论核心。徐通锵先生提出线性结构是现实的、表层的结构，而非线性结构是隐蔽的、深层的结构。非线性结构中的不平衡性是变异之源。结构的不平衡性可能在各个层面产生变异，而每一层面的变异又可能给其他层面或层面之间的结构带来不平衡性，因而语言的发展永远不可能堵塞结构不平衡性的漏洞。结构关联是指结构要素、结构子系统之间存在着的非线性的相互依存和制衡的关系。“依存”和“制衡”指各个结构项之间在运转和演变中的制约作用[①]。

（七）语言学的研究目标从“结构系统”转向“语言机制”

乔姆斯基生成语言学认为，语言学的研究目标不在于描写语言系统，而在于发现语言机制或语言能力。语言能力的发掘与智能机的制造成为当代国际语言学研究的共同目标。

（八）“语言静态同质说”被“语言有序异质论”所代替

弗斯认为，语言的异质性和非联系性，要比大多数人愿意承认的还要严重得多。拉波夫的变异理论认为，语言是一种有序异质系统。变异是有序之源，语言在有序的变异中不断改变结构。只有把语言放到社会环境中去，才能够发现语言变异的控制因素。

---

① 详见徐通锵《语言论》，东北师范大学出版社1997年。

(九)符号的"任意性原则"被"可论证性原则"所包容

1939年,本维尼斯特证明符号根本没有日内瓦学者所设想的那种任意性。1940年,巴利提出,符号的生命就存在于以内在必然的联想为基础的全约定和外部随意的联想为基础的全任意这两极之间。1954年以来,布达哥夫、斯米尔尼茨基、兹维金采夫对符号任意性原则进行了批评。1986年,游顺钊在《能指的起源》中认为,应该把"不知道"与"任意性的"分开。这两个概念正好反映了能指演变过程两端的特性,即源头"不清楚"而从现状来看则似乎是"任意性"。1986年,李葆嘉向全国首届青年现代汉语学术讨论会提交了《论语言符号的可论证性》。其中部分内容题为《论索绪尔语言符号任意性原则的失误和复归》(1994)发表,引发90年代语言观讨论。语言符号处于本原可论证性与演变任意性、任意性又被可论证性覆盖的矛盾中。

依据当代语言学的视界,语言研究可以分为描写、阐释和模拟三个层面。描写可分为静态与动态;阐释可分为机制和动力;模拟为语言形式化研究,其目标是智能机的制造。尽管描写是研究的基础,但描写并不是一切,应当扬弃的是"描写主义"而不是描写。尽管静态结构是科学的假设,但静态结构并不是一切,应当扬弃的是"结构主义"而不是结构。随着全面超越静态语言学的实现,近一个世纪来广泛影响了众多人文科学的结构主义,雄踞学术主流的时代一去不复返!随着语言和言语、内部和外部、静态和动态、形式和实体、描写和阐释的人为疆域的全面消解,语义研究和语形研究的合流成为世界性趋势,新世纪语言学旗帜上写着:"就人的言语活动、为人的目的而研究语言"。

## 第二节 语言学大师之谜和心理索绪尔

虽然经过戈德尔、恩格勒、毛罗以及斯留萨列娃对新发现的索绪尔的手稿札记等材料的研究,最终导致了"索绪尔的语言学思想的重新发现",但是仅仅停留在与"《教程》索绪尔"不完全相同的"手稿索绪尔"中显然不足。值得注意的并非《札记》中并不否定历时分析的论述,而是对历时研究深为不满的议论:"语言学家要研究的是语言态,他不需要理会导致目前语言态的历史事实,他应该把历时研究置于不顾。……历史的干预只能歪曲他的判断。"甚至措辞激烈地呐喊:"我们必须做出反应,抵制老学派的邪道,而这种反应的恰当的口号是:观察在今天的语言和日常的语言活动中所发生的情况……"为什么?! 曾为历史语言学家的索绪尔,为什么要吹响与历史语言学派决裂的号角? 为什么要振臂疾呼"抵制老学派的邪道"的战斗口号? 为什么要高扬静态语言学的反叛旗帜?

### 一、大师索绪尔身后留下的疑团

他的学生巴利,在《教程·第一版序》不无遗憾地写道:"他没有因此出一本书,凡特别有幸听过这门内容充实课程的人都深以为憾"。也许,读过《教程》的每一位学人,同样会和他的学生一样,为索绪尔生前把《教程》讲授提纲的草稿"随写随毁掉",没有加以整理,并进一步修订成为专著而深为惋惜。感慨之余,也许会发出一系列疑问:1.索绪尔为什么在讲完课就把讲稿扔进字纸篓,而不保存整理呢? 这是不是一种"怪癖"呢? 2.为什么没有以论文的形式来公开发表这些重要的理论呢? 这是不是由于他认为没有必要呢? 3.为什么生前没有尝试写一部全面阐述

这些理论的专著呢?这是不是他愿意且能够撰写,却出于谨慎或没有时间撰写呢?4.除了青年时代的两篇论文——《论印欧系语言的原始元音系统》和博士论文《梵语中绝对所有格的用法》,为什么没有再发表或留下其他重要论文呢?5.在日内瓦的几十年学术生涯,索绪尔是如何渡过的呢?

迄今为止,笔者没有看到学术界公布索绪尔公开发表其他论文的资料。甚至连哥本哈根学派的代表人物叶尔姆斯列夫在《语言的结构分析》中也深为困惑:“很难说,索绪尔的观点是如何在思想中具体形成的。”通过对索绪尔学术风格的探讨,通过对其理论直接渊源的寻求,我发现了一个“学术索绪尔”。在揭出索绪尔语言理论的三个直接来源和了解到索绪尔“不断革新”的治学风格的基础上,如果对活生生的作为人的索绪尔或“心理索绪尔”进行精神分析,兴许可以解开这个“索绪尔之谜”。

首先,索绪尔在博采名家和不断探索中给学生讲课,对于流动的思想来说,以别人的学术观点和理论体系为基础编写的讲授提纲草稿没有必要保存。其次,这些主要理论原则不是他的独创或首创,在他还是学生之前,博杜恩与克鲁舍夫斯基已经写成论文和专著发表了。在相互认识以后,博杜恩并且寄给了他一些。对别人已经发表的观点加以补正而敷衍成文,不是索绪尔这样一心要创立自己的理论的探索者愿意做的工作。再次,当时的欧洲还不像当代中国那样,盛行“兼收并蓄”(实质为抄袭或剽窃)别人的研究成果以编写自己署名的专著或教材。也许,他的学生既不了解博杜恩师生的研究——《教程》中提及众多学者的姓名,但是没有出现博杜恩师生的名字。到底是学生删除了博杜恩的姓名,还是老师没有提及博杜恩的姓名,无从详考——又不了解老师的烦恼,以致于在善意塑造“《教程》索绪尔”的同时留下了这些惋惜和一系列疑团。

## 二、中老年索绪尔的神经症缠绕

有证据表明，一心寻求变革的索绪尔陷入了深深的思考和长期的困惑。

1894 年 1 月 1 日，索绪尔致书他在巴黎的学生梅耶，谈到他不得已要交给编辑的一篇文章时，极言思虑与写作之焦虑："……可是我对这一切都厌倦了。厌倦的另一个原因是我，通常感到要在语言学的问题上，即使仅仅写上十行言之成理的文字也感到困难。长期以来，我一直在首先思索：如何把语言诸现象加以逻辑分类，又如何把研究语言诸现象的各种观点加以分类。我越来越认识到，要阐明语言学家所研究的是什么，需要做的工作量太大了……当前流行的术语极不妥当，有必要加以改进。为了改进这些术语，使我对语文学(历史比较语言学又叫比较语文学——李注)的兴趣越来越小，尽管我非常希望不要让我去概括语言的本质，这会使我违心地去著书立说。毫无兴趣和缺少热忱，去解释语言学中使用的术语为何，对我来说没有任何意义。然而，坦白地说，我只有写成这部书[①]，才可能在我停止的地方重新开始工作。"[②] 通过字里行间，可以看出中年索绪尔正处于极度不满、厌倦、焦虑，甚至陷入"失思"和"失写"的颓唐之中。

此时的索绪尔 38 岁，正当盛年的他可是竟对自己的研究一筹莫展，陷入厌倦和焦虑的神经症。也许，多年来的教学研究工作的殚精竭虑；也许，对理论语言学的期望值过高的困惑，造成

① 根据索绪尔对学生雷德林格的回答，大概指静态语言学著作。

② 见《索绪尔研究集刊》21 辑，1964 年出版。转引自许国璋《论索绪尔的突破精神》，《许国璋论语言》，外语教学和研究出版社 1991 年；又参阅了 J·卡勒《索绪尔》中译本，13 页，但做了文字上的些许微润色。

了他的神经衰弱；也许，遗传基因导致了他的心理疾病；也许，早年生活中的种种不快所形成的潜意识，促使他此时已经进入了中年心理危险期，并一直延续到晚年。由此，我试图揭出一个深层的"心理索绪尔"，也正是这个"索绪尔"使他焦虑不安，无法定下心来，建构完整周详的首创的语言学体系。

当他的学生雷德林格(Riedlinger)问他为什么不把普通语言学讲授的内容写出来时，晚年的索绪尔微露笑容，接着说："我没有给自己规定要写出静态语言学"①。面对难以接受这一回答的学生，索绪尔只是反复重申这项工作的困难。也可能，他早就预感到自己确实难以超越博杜恩创立的喀山学派的语言学理论高峰了。

## 三、青年索绪尔的首创权情结

有一个早年的事实，可能与索绪尔中晚年的心理疾病密切相关。

1976年，索绪尔进入莱比锡大学求学，经胡布施曼(G. Hubschman)教授介绍，了解到布鲁格曼(Karl Brugman，1849—1919)② 教授关于希腊语中的某些a由n变来的论文《印度日尔曼基础语中构成音节的响鼻音》。但是，索绪尔却对胡布施曼说："这一发现并没有特别的价值，也算不上是新的"。听者必定疑问：一个刚刚入学不久的新生何以出言如此不逊？其实，早在3年前，索绪尔已经发现了a＝n这一等式，但是他没有留下书面证据。卡勒的《索绪尔》中写道："可自己又否定了"，根据未详。布鲁格曼的发现无疑会给他增加自信。索绪尔感到，他的

① 戈德尔《稿本溯源》，30页。转引自《许国璋论语言》，106页。

② 著名历史比较语言学家，新语法学派的首领之一。

构想比布鲁格曼受到人们赞赏的假说毫不逊色，这可能就是促使他急于撰写《论印欧系语言的原始元音系统》的直接动机。1878 年索绪尔的论文发表，其中自然包括 1873 年他发现的响音内容，但他不仅只字未提自己的独立发现（所谓“首创权”），反而违心地写上“感谢布鲁格曼和奥斯特霍夫（Hermann Osthoff，1847—1907；新语法学派的主要成员）的著作，我们知道了响音 n 和 r”。

历史证明，这是一篇天才的论文。50 年以后赫梯语楔形文字的发现，印证了索绪尔对与 a 相关的、但与 e、o 无关的“共协半元音”原始印欧语喉音 h 的天才推断。文章发表以后，他的老师布鲁格曼只写了一篇礼节性的短评。但是，后来有人说索绪尔这一不借助外力独立完成的论文抄袭甚至剽窃了别人的成果，尤以奥斯特霍夫教授为甚。然而，引起新语法学派反感的深层原因，我以为并不在于 a = n 这一具体发现，而在于“乳臭未干”的出言不逊：“我不是在空想费解的理论问题，而是在寻求这一学科的真正的基础。没有这个基础，任何研究都是没有根据的、武断的、不能确定的”[1] ——触动了新语法学派头面人物的某根神经。这些流言蜚语，来自于他青少年时代一直推崇的新语法学派的师辈名流，使 21 岁血气方刚、雄心勃勃、崭露头角的索绪尔感到非常愤慨、非常苦恼、非常失望，由此形成他后来对“老学派的邪道”即新语法学派怒气冲天、出言不逊的潜意识压抑心理。

20 多年过去了，当德国学者斯特莱特贝格（W. Streitberg）写信给索绪尔谈到《论印欧系语言的原始元音系统》时，尽管事隔多年，然索绪尔仍难以释怀，为了澄清与新语法学派之间的关

[1] 见索绪尔《论印欧系语言的原始元音系统·序》，转引自卡勒《索绪尔》。

系,特别是为了洗刷"剽窃"的不白之冤,索绪尔还是写了《关于青少年时期和求学年代的回忆》[1],寄给斯特莱特贝格保存这份手稿,立此存照。《回忆》中,虽然他一再申明,他对首创权并不介意,但是恰恰据此可以发觉他对"剽窃"中伤的耿耿于怀。正是这青年时代的所谓"剽窃",在索绪尔心灵深处形成了"首创权情结"或"剽窃恐惧情结"(潜抑心理防卫机制)和"憎恶新语法学派情结"(升华心理防卫机制),以致于他后来呼吁"应该把历时研究置于不顾……历史的干预只能歪曲他的判断",以致于他几十年不写或极少写、更谈不上发表任何论著,惟恐因为观点的"不谋而合",而成为师辈指控他的"新证据"。

由此,一方面促使他"这位有造诣的新语法学家"力图冲破师辈所热衷的"老学派的邪道"或新语法学派的藩篱,创立新的语言学理论,以证明自己的独创性才能;另一方面,当他了解到博杜恩崭新的语言理论时,也许在赞赏之余不免顾影自怜。证明独创性的热望,燃起了他寻找更新的突破口熊熊之火,永无休止的探索和困惑,又致使他陷入了深深的焦虑。一筹莫展的他,"对这一切都厌倦了";忧心忡忡的他,不可能把普通语言学讲义编成专著。所有这些渴望、寻求、探索所带来的困惑和疑虑,导致他"仅仅写上十行言之成理的文字也有困难",只得以"我没有给自己规定要写出静态语言学"的微笑(合理化心理防卫机制)来掩饰心灵深处的难言之隐。

## 四、探索者索绪尔的人生悲怆

中年的索绪尔,后来有没有去做他在给梅耶的信中所说的"只有写成这部书"的工作呢?

---

[1] 约于1903年,索绪尔时46岁。布鲁格曼与奥斯特霍夫还在世。

非常遗憾，根本没有。厌倦焦虑的索绪尔却去以研究中世纪德国的传说文学为消遣，留下一些根本不想发表的笔记。晚年的他沉湎于研究拉丁诗人怎样在诗中隐藏专有名词的“字谜”之中，并且留下了大量“迷惑不解”的笔记。他在一封信中说：“在最重要的问题上我仍然迷惑不解。也就是说，对字谜这种现象或假设，应该如何解释呢？”①

这种字谜把字母分散在文章里，有时按原词的字母顺序出现，有时变换字母先后顺序或出现次数不等。依据精神分析法，“字谜”的成因可能是：因为有一个关键词在作者潜意识中徘徊，以语音联系而影响到其他词语的选择。问题不在于索绪尔的发现证明了“字母在潜意识中的坚持性”，而在于索绪尔沉湎于“字谜”发现的行为反映了他的什么“人谜”或精神状态？无论是对语言研究厌倦的一种排遣，还是以“钻牛角尖”的方式在逃避什么，这一切都表明他事实上对“在我停止的地方重新开始的工作”的“可能”渐渐丧失信心。而是以否定作用、转移作用这些心理防卫机制在减轻自己的心理痛苦，以保护自己。一个焦虑不安的人，通常会做出一些常人以为的“怪事”，难怪有一位评论家针对他研究“字谜”的情况说：“索绪尔有些神经失常”②，准确的说法是索绪尔长期患有身心疾病“神经症”。

在迷茫中的他，却始终没有发觉：除了共时与历时的对立，内部语言学之外的外部语言学是一条光明大道，20 世纪中叶兴起的边缘语言学证明了这一点。

1906 年，一位教授退休以后，学校让他负责普通语言学的教学。从那时起，他隔年授课（1907，1908—1909，1910—1911），

① 转引自卡勒《索绪尔》，卡勒引自 Jean Starobinskid 的《言外之意》。

② 转引自卡勒《索绪尔》，146 页。

可能又燃起建构静态语言学的宿愿。不幸的是,长期的探索思虑和身心失调,致使他1912年夏罹于沉疴。1913年,索绪尔带着探索者的困惑和遗憾在疾病中离开了人世,终年56岁……

然而,这并不否认索绪尔在探索和传播现代语言学理论过程中的创造性、系统化的发展,尽管以极端方式割断了历时和共时的关系,把共时放在绝对优先位置,但是纯粹的"静态语言学"无疑是他的创造。更要感谢他的学生所编行的《教程》,毋庸置疑,是"《教程》索绪尔"把国际语言学和众多人文科学推向一个新的历史时期。需要特别强调的是,"学术索绪尔"是一个淡泊名利的人,卡勒的《索绪尔》中有这样一段话:"虽然在老同事的要求下,他被授予勋级会骑士称号,可是这样的荣誉并没有能够把他留在巴黎"。索绪尔认为,他毕生的任务就是从事学术探索和培养新一代语言学家。也许,学生们的善意来自老师生前的关爱;然而学生不知,老师对他们的关爱正是来自自身青年时代所深受伤害的隐痛。

"手稿索绪尔"表明,他是一个思想活跃、富有洞察力的、不断进取的学者。正如博杜恩对语言学研究的兴趣是随着时间的推移而改变,并不刻意写出全面系统阐述"现代语言学"观点的著作一样,以开放的、流动的思维面对"封闭的"(其实是封闭与开放的统一)、"静态的"(其实是静态和动态的统一)语言系统,写出一本系统的完全首创的专著对于索绪尔也是勉为其难。

面对语言符号背后的庞大的认识系统,迄今为止,所有语言学家所做的一切研究还都是"划地为牢"、"盲人摸象"……。"人生之有涯,而知无涯",企图以个人有限的认知能力处理人类无限的认知系统,这种"学术集体无意识"导致了所有探索者焦虑、困惑和遗憾的人生悲剧,任何博大精深的学者也概莫能外。然而,正是知识的无限性和科学的可证伪性,求知和创造的冲动才

使一代又一代学者不无遗憾地舍弃人间的许多欢乐，心甘情愿地为悲凉而壮丽的科学事业贡献出有限的一生。

## 五、打开索绪尔的“黑匣子”

索绪尔的生平和学术风格，巴利的《教程·第一版序》和卡勒的《索绪尔》做了一些叙述，比如思想非常活跃，语言学思绪有如一条流动的河。然而，都没有阐明：是什么动力驱使原来醉心于历史语言学的索绪尔一心要冲破新语法学派的藩篱，追求创造新理论的梦幻？是什么原因致使中老年的索绪尔把《教程》讲稿随写随毁，生前没有写出静态语言学的专著？采取与索绪尔同时代的弗洛伊德所开创的精神分析法，通过剖析“心理索绪尔”，也许可以打开索绪尔的黑匣子。

许国璋在《论索绪尔的突破精神》[①]一文中自问自答：“一个有意思的问题是，一个仅是大学最优等毕业，写过一篇出色的论文（但这篇论文里的预言的证实还有待几十年以后一种 Hittite 文物的出土），还在大学教书但是尚未有专著问世的索绪尔，何以敢于这样富有自信地纵横议论呢？我想，我们不妨把这种自信叫做哲学的自信！”在《从两本书看索绪尔的语言哲学》[②]一文中又修改了这一看法，而认为是索绪尔多年来的历史比较研究、多年来在大学的讲授，深知普通语言学的其中底细，晚年才试图创立新的理论。许国璋之所以提出这种根本不存在的“哲学的自信”的臆断和“深知其中底细”的揣测，是因为不了解索绪尔语言理论的直接来源和内心苦恼。因此，他虽然引用了索绪尔信中的“厌倦”，但却未能体察探索者的“悲怆”。

如果克鲁舍夫斯基不是英年早逝，如果博杜恩的学生谢尔

①② 《许国璋论语言》，外语教学和研究出版社 1991 年。

巴们将老师的遗著编成全面阐释现代语言学的专著,那么现代语言学鼻祖的桂冠也许不会落在索绪尔的头上。然而,“学术索绪尔”绝不是一个欺世盗名的人,以他“随写随毁掉”的习惯,生前根本不可能想到:在他死后,由于担心老师所讲授的内容或学术思想失传,出于学生们的善意,“《教程》索绪尔”才把他推上20世纪现代语言学“奠基者”的宝座;更没有想到与他的英名密不可分的“结构主义”会对语言学和众多人文学科产生如此巨大的影响。因此,只有分析了心理索绪尔,才最终打开了索绪尔的“黑匣子”;只有了解了表层结构和深层结构四位一体的索绪尔,才真正读懂了作为人的索绪尔及其《教程》。

当我们了解到这些,再来看卡勒的《索绪尔》(引文中的问号是我所加):“索绪尔一生过着平静的生活(?),……他没有经历过什么思想危机(?)或决定性的思想转变时刻(?),……由于他十分谦虚(?),我们难于追溯他早期思想的形成过程(?)。他未能完成什么主要著作,这似乎可以典型地反映出其矛盾的经历(?)”;“索绪尔的论文受到多方好评(?)……索绪尔在法国似乎感到不愉快(?)”;“在日内瓦……他结了婚,后来有了两个儿子。他很少出游,好象在一个体面而又偏僻的地方隐居起来(?)。他动笔愈来愈少,后来感到写作十分痛苦,更懒于动笔了(?)”。

这本写于1976年、目前国际上关于索绪尔的惟一的一本评传,除了对上述疑问没有追根穷源,在《索绪尔理论的地位》一章仅仅将索绪尔与同时代的德克海姆、弗洛伊德平行比较,未免浅尝辄止。但要考虑到,作为文艺批评家而非语言学家的卡勒能够写出这样的第一本《索绪尔》评传已经颇具开创性。不过,兴许应该承认:与伦敦大学校长夸克(R. Quirk)所言这本书“对索

绪尔的理论作了出色的介绍"[1] 的评价相比，中国语言学家许国璋"此书解释索绪尔本人的学说并无独到之处"的评论较为恰如其分。

[1] 张景智《索绪尔·译者序》

# 第五章　汉语史理论模式与东亚南洋语言文化圈

## 第一节　论汉语史研究的理论模式

### 一、汉语史研究理论模式述评

汉语史研究中的每一具体项目的操作都与特定的汉语史观密切相关，尽管一些研究者并非明确意识到这一点。那种不适当地夸大语言材料的自明性而漠视理论模式建构的倾向，除了表明已经接受了某种模式的支配以外，还隐含着理论思维的薄弱。语言史的研究脱离理论模式，其结果难免只见树木而不见森林，流于堆砌而惑于贯通，难以揭示语言演变的动力、轨迹和机制。

20世纪上半叶，瑞典汉学家高本汉（Bernhard Karlgren，1889—1978）建立了直线型理论模式；60年代中期，美国普林斯顿学派建立了立足方言的逆推型理论模式；70年代初期，美籍华人张琨建立了南北汉语的差异型理论模式[1]；70年代后期，日

[1] 白丁《严学宭先生谈汉语史研究动态》，《语言学通讯》（油印本）1985年总第5期。

本桥本万太郎建立了农耕语言的推移型理论模式。虽然其中仅有桥本的一种模式，涵盖基本词汇、语法结构和语音结构三个子系统的研究，其余三种模式主要是针对语音历史演变的，但是建立在某种语言史观之上的理论模式，即使从某一语言要素或子系统切入，也必然会以种种方式牵涉到其他语言要素或子系统而牵一发动全身。就这一意义来说，汉语语音史研究的理论模式也就隐含了汉语史研究的理论模式。

20 世纪，在中国占主导地位的是高本汉建立的直线型模式。早年的高本汉（1915—1926）在成名作 *Etudes sur la Phonologie Chinoise*[1] 中表达了自己的研究旨趣：

> 1. 把中国古音拟测出来，作系统的现代方言研究的起点；2. 把中国方言的语音作一个完全描写的说明之后；3. 然后可以用音韵学的研究指明现代方言是怎样从古音演变出来的。（《中国音韵学研究》13 页）

中年的高本汉（1940）完成了以字典形式反映其上古音构拟体系的《汉文典》（1957 年出修订本）。晚年的高本汉（1954）出版了他毕生研究中国古音的总结之作 *Copendium of Phonetics in Ancient and Archaic Chinese*[2]，论述了汉语音韵构拟的材料、方法和结论，表述了他研究汉语语音史的基本思路：

> 我们用“中古汉语”（Ancient Chinese）指代《切韵》纂集

① 《中国音韵学研究》（赵元任等译），商务印书馆 1940 年初版、1948 年再版，1995 年影印

② 《中上古汉语音韵纲要》（聂鸿音译），山东出版社 1989 年。

的公元600年左右的语言，这实质上就是陕西长安方言。这一方言在唐朝成为一种共通语(Koine)，除沿海的福建以外，全国各州县的知识界人士都说这种语言。另外，用"上古汉语"(Archaic Chinese)表示西周时代(从公元前1028年起)河南地区的语言。

有理由说，各地的底层民众在很大程度上保存了本地区的方言俗语，许多地方的土话里仍然可以分辨出中古方言的痕迹。但共通语却极为广泛地传播开来，并且被从最高级官吏到中下阶层的绝大多数人所接受了，从而成为几乎是全部现代方言的母语(福建与毗邻地区的闽方言除外)。《切韵》韵类和各个现代方言韵类之间的严整对应确凿地表明，《切韵》描写了一种实际存在过的、单一的语言，而不是像许多近来的学者所论述的那样，《切韵》是一种人造的、由各方言中的参差成分所构成的折衷混合的语言。

(《中上古汉语音韵纲要》2页)

这就是高本汉以《切韵》为中枢，上推上古汉语，下接现代方言的直线型理论模式的表述。这一模式的预设是：1.《切韵》音系为隋唐陕西长安方言；2.在唐朝成为一种共通语；3.几乎是全部(除闽语)现代方言的母语；4.上古西周汉语是河南语言；5.《诗经》韵脚音系和谐声音系代表上古西周河南语言；6.隋唐陕西长安方言由上古西周河南语言演变而来。

所谓中国古音学界的高本汉传统，包括具体研究方法和宏观模式这两个方面，而后者可以称之为"高本汉模式"，虽然研究古音的历史比较法可取，但阐释音系演变的宏观模式却违背中国语言史实。深受欧洲19世纪进化论思潮熏陶的高本汉，认为汉语发展从上古到古到现代方言是直线的，由周秦古音演变为

《切韵》音系，又分化为现代汉语各方言直线型模式以中古构拟为枢纽，上推周秦、下接方言而出现了汉语演变史的那种"蛇吞青蛙的隆腹现象"①。其症结是把《切韵》音系曲解为单一音系，一方面认为《切韵》是中古长安话，另一方面又认为现有汉语方言都是《切韵》子语，因而陷入自相矛盾。由《切韵》上溯至《诗经》，则形成从上古到中古由简而繁的分化，由《切韵》下推方言，则形成了从中古到现代由繁而简的合并。这种简→繁→简的橄榄状不符合语言演变的状态。直线型模式的失误有三个方面：1. 抽去了语言的横向空间差异，仅仅以一维的纵向时间差异来贯穿纷繁杂芜的语言演变，2. 以建立在印欧语基础之上的谱系树理论硬性框范中国语，用语言分化论处理中古汉语与现代方言的关系。3. 撇开汉语的历史文化背景，并不了解汉人对汉语音韵的心理感受，高本汉的局限并不在于"他的小学根基远远不能与章太炎黄侃等人相比"②，而在于他不具备汉人文化心理。由于囿于历史语言学"同一条件下的同一音位，在同一时间和空间只能产生同一音变"的封闭系统性原则，而不了解汉语语音演变的复杂性，这就导致高本汉在古音构拟中一味地追求"分"；由于囿于上古汉语语音和中古汉语语音的直线流变，而不了解汉语语音演变的错综性，这也导致高本汉在古音构拟中一味地追求"分"；由于务必详细阐释上古汉语语音到中古《切韵》音系的膨胀性发展，必然导致高本汉采取音素法而避开音位法。并非"高本汉虽然自觉地摆脱了清代小学的桎梏，但又不自觉地

---

① ［美］张琨《汉语音韵史论文集》（张贤豹译），华中工学院出版社 1987 年。在这之前黄淬伯先生早已注意到这一结果，见《论切韵音系并批判高本汉的观点》，《南京大学学报》1957 年 2 期

② 聂鸿音《中上古汉语音韵纲要·前言》。

陷入了历史语言学的樊笼"[1],而是高本汉就是在以印欧历史语言学眼光审视与印欧语发展模式不同的中国语音史,直线型模式就是用谱系树理论硬性框范东亚语言,用语言单源分化论处理中古汉语与近代方言关系的结果。直线型语史观的基础是"一元扩散论"的文化史观,而这一模式隐含的汉文化中心论满足了潜在的大汉民族情结,并且因简单化而便于研究时操作,而得到了中国语言学界的普遍认同,尽管有所调整,但基本模式未变。

高本汉模式在20世纪60年代以后受到批评,新的模式随之产生。美国加州伯克莱大学教授、美籍华人学者张琨,从时空二维研究汉语语音史,强调南北方言差异,建立了差异型理论模式。他的主要成果有《〈切韵〉的综合性质》、《论中古音与〈切韵〉之关系》和《汉语音韵史中的方言差异》等。其高足张贤豹先生以为:"张琨教授夫妇(1972)合著的《古汉语韵母系统与〈切韵〉》是20世纪初高本汉《中国音韵学研究》问世以来最值得重视的汉语音韵史著作。1972年是汉语音韵学的分水岭,1972年以前是高本汉时代,1972年以后是张琨时代。"[2] 张琨重申《切韵》是一部具有综合性质的韵书,代表的是公元601年以前若干百年不同地区的方言,由此可以投影出原始汉语系统。他批判道,把原始汉语设想为一个语言,后来才分裂为方言群——如先分裂成原始吴语、原始闽语等等,然后再分裂为各个方言——这是荒谬的假设。方言差异无疑是自古而然的现象,早期的汉语方言必定比今天更复杂。由于语言接触机会增多,方言才越来越加感受到标准语统一的影响力。因此,所谓汉语的原始系统,不是

---

① 聂鸿音《中上古汉语音韵纲要·前言》。

② 张贤豹《汉语音韵史论文集·代译序》。

一个历史上的语言，而是一个假想的对立系统，这是为了用最简单、最合语言实际的办法来解释已知的历史文献上的记录。不仅现在的方言差异是南北方言平行发展的结果，而且文学语言基础也有南北之异。周秦、汉代的文学语言基础是北方方言，而齐梁时代是南方方言，唐宋以降复以北方方言为基础。与高本汉的时间一维语史观相比，张琨的南北差异、方言分歧的时空二维观较切合汉语实际。但一些问题尚须深入探讨。何为古代北方方言，何为古代南方方言？南方方言是如何形成的，北方方言又如何变化？齐梁的南方方言（文学语言基础）是江南吴语，还是南迁中原汉语的南方化？唐宋元与明清的文学语言基础——北方方言是否有地区之异？中原正音、江淮官话、燕代方言与唐宋以降的文学语言基础各有什么关系？先秦的文学语言基础是北方方言，是否意味着《诗经》音系与《楚辞》音系同为一种？如果以为把原始汉语设想为一个语言是荒谬的假设，又如何认定早期的一些语言是汉语方言？原始汉语又从何而来？原始闽语、吴越语、楚湘语等与原始华夏汉语是同一语言还是不同语言？中国古代的"方言"概念与西欧的"方言"概念有何异同？凡此种种，只有结合考古文化系统及相关史料，只有把华夏汉语的形成与延伸置于农耕文明与游牧文明持久冲突交融的历史背景上审视，才有希望解决。差异型模式虽然较直线型模式有所进步，但仍然未摆脱狭隘的文明一元分化论，对中华大地上区域文明的复杂性把握不足，因此，在文化史观上与高本汉尚无根本性区别。

20世纪60年代中期，在美国普林斯顿大学中国语言学计划研究中的一些学者，批评了高本汉的研究方法，主张以现代方言为研究古代汉语的基础，而不必参考古文献记录的证据，被称为Princeton hypothesis（普林斯顿假说）。他们的操作程序是，

首先为每一种单独的方言群分别构拟出它们的原始语,如原始官话、原始粤语、原始吴语和原始闽语等,由此再构拟一个原始汉语。构拟出来的原始语,不像一个统一的一致的语言,而是带有内部变异的若干种历时系统。这种方式姑且称之为方言逆推模式。这一派主要学者及研究成果有:华盛顿大学罗杰瑞(Jerry L. Norman)的《原始闽语的声母》(1974)和《原始闽语的韵母》(1981),耶鲁大学司徒修(Hugh M. Stimson)的《汉语原始北方话的研究:北京话的阴平调》(1969),乔治亚州立大学贝乐得(William L. Ballard)的《原始汉语:塞擦音初探》(1968)和《原始湘语及其他》(1970),华盛顿大学余霭芹(Anne Y. Hashimoto)的《原始粤语的辅音和复辅音》(1970),乔治城大学杨福绵的《原始汉语的前缀 * s-》(1975),还有欧柯诺(Kevin A. O'connor)的《原始客家话》(1976)等。普林斯顿学派这一名词已经成为一个语言学术语,即"有兴趣以现代方言资料为基础进行构拟的汉学家"[①]。

普林斯顿假说牵涉到如下问题:1.现代方言区内部的差异度。各大方言区内次方言或土语群之间可用于比较的差异度各不相同。北方话(不包括晋语、江淮语)内部较为一致,可比性不高。闽语有"十里不同音"之说,内部差异性较大,可比性较高。语言内部的差异度决定可比性,而可比性与构拟结果的复原度有着一定的制约关系。2.现代方言的存古度。依据现代方言为基础构拟出来的早期语言,约当于什么年代的汉语面貌。如抛开文献记录,一味依赖活语言构拟,其语言年代则无从确定。若以传统的中古音系为参照物,各方言的存古度是不同的。有证

---

① 殷方《1965-1979年国外汉语音韵学研究述评》,《汉字汉语学术研讨会论文集》(下),吉林教育出版社,1991年。

据表明，南方方言原来是非汉语言，由于汉文化与汉语的扩展才使当地土著语演变为汉语方言，因此这些方言中有可能残存土著语的底层成分。撇开这些非汉底层成分，由于影响的年代不同及方言演变的差异，闽语、粤语、吴语、湘语、客家话、徽语、赣语、江淮语、晋语、北方话各自的存古度大致呈递减状态。若以现代闽语为基础构拟出来的早期闽语可能约当于中古以前的汉语言，若以现代江淮语为基础构拟出来的早期语言可能约当于宋元明之间的汉语言，但若以北方话为基础构拟出来的早期"豫燕语"则只能约当于元明以后的汉语言。可见，依据现代各方言构拟出来的早期方言并不处于同一语史平面上，也就为逆推原始汉语带来极大困难。3. 南北语言的同源性。根据考古文化，中华大地上的新石器时代没有单一的原始母文化，这就表明没有相应的单一的原始母语，人种体质类型的差异及演变、历史文献的记载也支持着这一观点。与西欧古希腊方言是仅在某个别音素上有所差别的同一语种不同，中国周秦汉时代的"方言"即邦言，是不同的语言。它们之间的差别不仅是语音的差别，而且存在词汇和语法结构的不同，因而中国最早的方言学著作扬雄的《方言》表现为词语的网罗，而"转语"则是不同邦言之间借词引起的结果。现代汉语的几大方言的形成是秦汉以后华夏文明与语言对周边不同的文明与语言逐步扩散且同化程度各别的结果；而在这一过程中，华夏汉语也受到周边语言的浸染而发生了剧烈的嬗变。

普林斯顿学派试图以不平行的早期方言为基础，逆推出原始汉语，其结果还是个谜。虽然他们使用的材料及操作方法与高本汉大不相同，但立论的语史观还是相当一致的。因而高本汉曾经有过利用方言直接构拟早期汉语的设想，也就不足为奇了。显而易见，利用现代方言为基础构拟早期汉语的逆向推溯

模式,是对一味依赖文献、仅将方言作为构拟音值参考的高本汉传统的反动与挑战,但是全然抛开历史文献,将汉语的演变发展与没有文献的语言的演变发展等同视之,不仅不能恰当地重建历史汉语,而且构拟会不着边际。

高本汉与普林斯顿模式都是以"西洋镜"观察汉语史,难免游移于中国文化史背景之外。张琨模式已注意到结合中国政治与文化中心的转移来考察汉文学语言基础的变异,但对汉民族主体的形成与华夏文明的生成与延伸背景尚未能予以足够的重视。桥本万太郎(1932—1987)从文明类型入手,将东亚大陆语言确定为与印欧"牧畜民型语言"不同的"农耕民型语言"。在《语言地理类型学》[1]里建立了与西方语言谱系树说不同的推移理论模式。他认为,既然东亚大陆语言的现状在地理上表现出的类型推移是语言历史演变各层次的投影,从而便可以把"横"的区域考察与"纵"的历史比较有机地结合起来,去追寻语言发展的踪迹。桥本审察汉语发展的视域已不限于汉语、汉语方言,而扩展至周边语言,构画出东亚大陆语言推移结构连续体。他立足于东亚大陆的民族移动和社会变迁,一直是从北向南展开的脉络,而中原地带经历过北方游牧民族不知多少次的反复入侵,过去的十个世纪的大半时间都处在北方阿尔泰诸民族的控制下,提出了北方汉语阿尔泰化与汉语同化南方语言的论断。推移模式不仅针对语音结构,而且涵盖句法结构与基本词汇,因此是一个较为完整的语言史研究模型。桥本实际上运用了文化圈理论,语言结构连续体区域推移状态与这一文化圈内文明传播方式互为表里。

1 [日]桥本万太郎《语言地理类型学》(1977,余志鸿译),北京大学出版社1985年。

但是，桥本的观点受到大陆语言学界一些人的冷淡与抵制，据说主要是感情上接受不了“汉语是南北少数民族语言挤压形成”和汉语方言原是不同语言的结论。桥本理论草创，未能进一步细密，大陆一些学者又未加以深究，其中“汉语南北挤压形成论”即是一种误解，与桥本所论“北方汉语阿尔泰化并同化南方语言”不相切合。至于南方方言原与汉语是不同语言，这在先秦文献中有“五方之民，言语不通，嗜欲不同”的明确记载。西汉扬雄的《輶轩使者绝代语释别国方言》，“别国”二字已明其所录方言为邦国之言即不同语言，因此朝鲜语也收列其中。究其原因，国内语言学界受欧洲历史比较语言学影响太深，以为谱系树理论放之四海而皆准，又未及对东西方语言的文化模式差异加以探索，实际上是固守“高本”而拒认“桥本”。此外，大陆语言学界存在着十分严重的轻视宏观理论的倾向，满足于引进国外理论而缺少跨学科的理论研究。语言学家追求语言研究的“纯正”，对文化史的研究不甚关心，对史前文化了解很少，更不用说运用文化圈、文化区域的理论，虽然不满意桥本模式，可又未提出独到的理论模式。值得提及的是，桥本在写《语言地理类型学》之前，并没有访问过中国大陆。只缘身在此山中，不识庐山真面目，使人感到莫名的悲哀。桥本理论并非没有缺陷，推移论的基础“黄河文明中心一元扩展单动说”与 80 年代考古文化研究的成果大相径庭，但书成于 70 年代尚可理解。然而正是这一“黄河文明中心说”，又与中国语言学界奉行的高本汉传统赖以成立的文化史观有相通之处。

## 二、原始华夏语三元混成发生论

推移方式应是已经具有强有力的某个文明中心以后出现的态势。在中心形成过程中，层叠性混成才是这一中心区域语言

发展的态势。原始华夏语通行于黄河中游河洛地区,这一区域,北与阿尔泰胡狄,西与藏缅氐羌,东南与南岛夷越毗邻。华夏语与这些周边语言存在着千丝万缕的联系。换而言之,河洛地区成了阿尔泰、藏缅、南岛三种文化与语言的交汇之处。历史语言学家可以从不同角度分别探讨华夏语与这三种毗邻语言中的任何一种,在发生学上的同源或接触中的渗透关系。

依据汉语与藏语同质的部分,可以认为汉语和藏语同源,主张此说的西方学者有西门(W. Simon)等,中国学者有王静如、俞敏等,其代表作如俞敏的《汉藏两族人和话同源探索》。依据汉语与南岛语同质的部分,可以认为汉语与南岛语同源。20世纪初期,一些外国学者孔好古(A. Conrady)、吴克德(K. Wulff)等,通过词汇对比证明在澳斯特利语和汉藏语之间有发生学关系,试图提出一个汉藏泰澳超级语系重新解释东亚及南太平洋诸岛语言的发生学关系问题,并指出南岛语和汉泰语之间的关系比汉泰语和藏缅语的关系更密切些。与之同时,还有许多艺术史家、人类学家和考古学家,为了说明南岛文化与中国北部和中部早期文化的一致性而做了大量论证。沙加尔在第23届国际汉藏语言学会议上发表了《汉语南岛语同源论》,邢公畹同意沙加尔的观点[①]。依据汉语和阿尔泰语同质关联的部分,或认为汉语与阿尔泰语有发生学关系。竟成认为,汉语与阿尔泰语之间的关系可能反映了语系分化前这一地域的状况[②]。

20世纪70年代以来的中国考古研究成果表明,中华大地在新石器时代存在着三大考古文化系统,即青莲岗文化系统、仰

---

① 邢公畹《关于汉语南岛语的发生学关系问题》等,《民族语文》1991年第3、4、5期。

② 竟成《汉语史研究的新思路》,《现代语言学》1992年第21期。

韶文化系统与北方细石器文化系统。这三种文化创造者的体质特征，青莲岗文化为蒙古人种南亚型；仰韶文化为蒙古人种东亚、南亚型并存，青铜时代为东亚型；北方细石器文化为蒙古人种北亚、东亚型并存，晚期趋于混合[①]。这三种文化创造者的语言盖分别为：太古夷越语或原始南岛语、太古氐羌语或原始藏缅语和太古胡狄语或原始阿尔泰语。与人种体质特征类型的混合相平行，语言的混成融合也不可避免。

与三大考古文化系统相互证发的是古史传说中的三大氏族系统：兴于江淮而发展至黄河下游的伏羲氏太皓为夷越—苗蛮之始祖；兴于渭水而发展至黄河中流的神农氏炎帝为氐羌之始祖；兴于北方草原而后南下，直至江汉的轩辕氏黄帝为胡狄之始祖。部落战争促成了华夏族的形成。共工蚩尤战为羌夷之战，西北与东南文化交流。共工败，炎帝"乃说于黄帝"，炎黄结成联盟是氐羌与胡狄文化交流。黄帝擒杀蚩尤于冀州之野，是北方游牧民族打败东南农耕民族。黄帝与炎帝战于阪泉，"三战，然后得其志"，是北方游牧民族战胜西北农耕民族。黄帝乘势进居江汉，"五十二战而天下咸服"。黄帝的胜利奠定了华夏族的基础，因此被后世奉为"人文之初"。游牧民族强悍，征战常胜，然农耕文明先进，因此胜利了的统治者不得不被战败者的先进文化所同化。经颛顼、帝喾、尧、舜、禹的统治，华夏族在中原地区逐步形成，同时原始夷越语、氐羌语和胡狄语也逐步混合而成原始华夏语。而没有卷进融合漩涡的周边初民，或留居旧土，或四处迁徙。西北的远至青藏高原，朔方的远至大漠以北，东南的越

① 石兴邦《中国新石器时代考古文化体系及其有关问题》，《亚洲文明论丛》，四川人民出版社1986年；格勒《中华大地上的三大考古文化系统和民族系统》，《新华文摘》1988年第3期。

过长江，直至南太平洋诸岛。

胡狄系夏人虽未传文献但中原居民称"夏"始于此。夷越系商语的顺行结构 NA 为南方语言之特色。氐羌系周人来自西北，在其统治的几百年间使华夏语"氐羌化"，形成了以王畿方言为基础的雅言。周人自称有夏、诸夏，所谓雅言即夏言。没有同化的北狄、西戎、东夷、南蛮，则被诸夏视为"夷"，于是有夷夏之别。在氏族或民族融合中得以融形成并发展的华夏语，不可能是纯粹的而不是混成的，只是古代华夏语言的异质性与聚合性被传至后世的表意文字掩盖了。

不管是汉—藏同源说、汉—南同源说，还是汉—阿同源说，都是从不同角度，依据各自选择的语料、人种学与文化史方面的证据，分别论证了汉语与其他语言的同源关系。原始华夏民族的形成是史前民族或种族融合的结果，与之同步发展的华夏语不可能不具有与其被收融合语的对应之处。与其说汉语仅仅与某一语言有同源发生关系，不如说在发展的长河中，后来被称为"华夏汉语"的这种语言与毗邻三大语言的祖语都先后出现过渗透、换用与混成关系。换而言之，原始华夏语就是一种多元性的层叠性混合语[1]。混合语既是语言混合的结果，又是文化融合的现象。

在华夏语的混成发生过程中，语言的渗透、换用、融合相当错综复杂，引起了音节结构的变化以至语言类型的转变。现存中国最古文献多为周秦时代所撰，因为周人是氐羌与胡狄混血，所以周代文献《尚书》《诗经》中的语言既与藏语有同源之处，双音节词中又具有与阿尔泰语元音和谐相似的特征。因为周人因

---

① 李葆嘉《试论原始华夏语的历史背景》，《语言学通讯》1990 年第 1—2 期合刊。

袭了夷越系的殷商文化，而夷越南蛮又未能一下子同化，所以在古代汉语中仍然可以发现大量的与南岛语对应的词项。

## 三、秦汉以后汉语推移发展论

周联邦所在的黄河—长江流域，上古语言的分布呈现三字型格局：中部为华夏语……东夷语；南部为原始吴越语……蛮楚语……巴蜀语；北部为辽东语……幽燕语……三晋语。联邦的北境是阿尔泰语，西鄙是藏缅语，东南方是百越—苗蛮语的广大区域。西汉扬雄的《方言》不是 dialectology 专著，而是一部用汉字记录了汉语和周边民族词语的比较词汇学著作。"方"邦也，秦汉方言不同于古代希腊语中口音稍别的 dialect，而是邦国区域性的不同语言。秦汉以后所逐步形成的汉语"方言"，是一个以可追溯的文化渊源为内涵的人文历史地理语言学概念，汉语方言是历史上的汉人移民周边化和民族语言汉语化的产物。以永嘉之乱、安史之乱（包括藩镇割据和黄巢动乱）和靖康之难三大动乱为高潮，北方游牧民族进踞中原，引发了秦汉以后人口迁徙的一次次浪潮。在促使中原汉语分化和周边语言换用的同时，彻底改变了中国境内的语言分布，在造成了汉语主导语言的基础音系历史性变迁的同时，最终导致了近代汉语方言格局的形成。

第一次人口迁移高潮促成了汉语从上古到中古的突变。肇始于永嘉之乱的中原居民迁徙的结果，突出表现为中原汉语分裂为北留中原汉语和南迁中原汉语，打破了中国境内汉魏以来的语言分布格局，促成了语言运动的新的趋势。1. 沿江中原汉语带的出现，江淮移民总数多达 43 万，在吴楚语区的北境形成一条东起海滨、西达湖鄂的江淮语带。2. 中原语言南流走廊的形成，中原居民由江淮迁皖南或经鄱赣平原沿江而下，连成一片

的吴楚语从中隔断,逐步形成了江东语、徽语、赣语对吴语的弧形包围。3.洞庭北部出现了西南话的雏形,从秦雍沿汉水流域到达洞庭流域的居民成为西南话的祖先。4.浙江沿海越族的连锁性南迁,中原汉人到达江南以后,引发了浙江沿海百越的连锁性南迁,闽粤境内百越语逐步演变为闽语。5.汉人的南迁迫使南方民族的迁徙和语言的演变,连接长江与珠江流域的皖赣中原语言文化南流走廊的出现和沿海越人的南迁,波及越蛮俚僚等族的迁徙和语言演变。6.幽燕语和三晋语的汉语化运动,中州士女或迁往辽西或西北,推进幽燕和三晋的汉语化运动。7.北留中原汉语的阿尔泰化,战乱和流迁致使中原人口急剧减少,内迁北民超过200万,北族的语言换用引发了北留汉语的阿尔泰化。此为秦汉以来中国境内语言一大巨变。"北方"首次作为大语言区出现在东晋郭璞《方言注》中,《音辞篇》记载了南北语言的特征分析。安史之乱、藩镇割据和黄巢动乱,一再迫使迁徙到长江中游北岸的中原居民继续南迁。大量居民进入江西,促使赣客语成型并逐步南进。五代十国,突厥人逐鹿中原,促使中原汉语和北族语言的交融。此为中国境内语言又一巨变。肇始于靖康之难而绵延于南宋时期的人口迁移高潮,导致了汉语从中古向近古的彻底转变。中原之音由此四分五裂,有掳迁蓟北为华北方言底层之燕代方言,有迁往浙东为吴语同化之临安宋音,有迁往江淮保存宋音之明清官话,亦有滞留中原但难免阿尔泰化的汴洛方言。此为中国境内语言第三次巨变。

在这一历史性变迁中,永嘉以后的中原语言文化南流的涌动,致使连成一片的广大吴楚语被纵横相交的T字型语带所分割。在中原汉语覆盖土著语言的T字型格局中,地处江淮鄂东的北部吴越—南楚语,受中原语言覆盖形成江淮语,横向连成中原沿江语言带;江南的吴越—南楚语被中原语言南流走廊分割,

纵向形成了徽语、赣客语中南语言带。T字型的东西两翼也发生着变化：太湖流域的吴越语受中原语言渗透为中古吴语；湖湘流域的的南楚语受中原语言影响演变成中古湘语。中原居民的迁入又引起吴越土著南迁瓯闽。T字型最南部是中原移民进入岭南而形成粤语。东西两翼语言的变化波及到百越苗蛮，促使其语言结构逐步转型。显而易见，南迁中原汉语的纵横T字型语带，奠定了汉语方言格局形成的历史基础，决定了中原汉语的持续流向。

汉语在胡人进踞、士民南迁与民族融合中"北杂夷虏，南浸吴越"。阿尔泰化是北族在"换用中原汉语"的过程中以阿尔泰语底子异化了汉语；夷越化是中原汉语对百越语施加的"汉语化进程"而残留了百越底层。惟独地域上兼接南北的江淮语言带，相对保留了中原正音，其中古源头与南方汉语相似，其近古源头与北方汉语密切，从而呈现出北方汉语……中部汉语……南方汉语的连续体推移状态。

在阿尔泰居民大规模换用汉语的基础上逐步形成了北方汉语。阿尔泰居民大规模换用汉语可追溯到北魏，辽宋以降汉语阿尔泰化的人口日益扩大。契丹、女真、满洲的语言换用及蒙古的影响，是促使北方汉语迅速变化的根本动因。随着明清的大规模移民，北方汉语占据了西南、西北和东北，除了相对闭塞的山西一带存在具有特色的晋语。另外，尽管毗邻冀豫的上古齐鲁方言华夏化后卷入中原汉语中古后的演变旋涡，但地处海隅的胶东方言具有复辅音、中古通摄字今读-m尾、声调发育不足等，可能残存着东夷的孑遗。在阿尔泰人换用汉语或汉语阿尔泰化的互动中，大北方汉语迅速变化且地区不断扩展，最终导致了北方汉语带的连成一片。这一语言带的西部漫长边缘则形成了汉语和西北、西南民族语交汇的走廊。

在汉语对南方民族语覆盖、渗透、混合的基础上逐步形成了歧异显著的南方汉语。较大规模的中原居民入闽是永嘉之乱以后,从六朝到唐代,中原人口的陆续迁入致使福建人口增长10倍以上。唐时"福佬"的名称可作为闽地民系形成的确证,北宋太宗时"闽语"已见于文献。根据唇音合舌音的特点,或认为闽语承传了上古汉语的声母,但如注意到壮侗语一般没有轻唇、舌上,苗瑶语也没有舌上但已分出轻唇(腊乙坪苗语无轻唇),印尼、朝鲜、日本语也无轻唇音,那么闽语的重唇、舌头,到底是唐代以前南迁汉语的历史层面还是闽越语的底层值得探讨。福建地处山区而交通不便,早期汉语化程度不高,后期接受北方汉语的影响不及粤语,因此保留的百越语底层成分较多,语言年代较古老,以致于高本汉不得不另眼看待。

汉人先秦就进入粤地,但直至北宋广东总户口数突增为唐代的2.7倍,近古粤语的基础才最终稳定。粤语区域地处珠江流域,居民越过南岭可望抵达,难免受到历代中原汉语的陆续影响,其语言特征约当于唐宋中原汉语层面,并且存在百越语的早期渗透和后来的持续影响。在广西,与主要由于宋代军事移民奠定基础而形成的平话不同,桂北土话源于唐五代零星陆续迁入、方言来源歧异的北方汉语的南迁驳杂土语群。土语群和平话既不具有历史上的同一移民源,又不具备演变中的环境一致性,因此二者不属于同一方言。

江南太湖流域古称"包孕吴越",《越人歌》反映了古吴越语与侗台语同源。泰伯带去的先周语,因人口比例太低而影响不大。直至两晋之交的移民完全覆盖了古吴北境的江淮并深入太湖流域,中原汉语和土著越语的融合才形成中古吴语。靖康之难的北方移民,不仅在浙北造成杭州方言岛而且推进到浙南,中古吴语逐步演变为近古吴语。与古吴越毗邻的是荆楚,楚人先

周时由中原进入鄂湘，和苗蛮混合而成蛮楚语，至秦汉时演变成南楚语。《方言》中提及的“南楚以南”的语言，则是指苗蛮—百越语。晋以前中原人取笑楚人语作“哑哑声”。永嘉中原居民的南迁，致使南楚语演变为古湘语。近古以后受西南方言的渗透，分化出浊音消失的新湘语。面对新老湘语受近古汉语化程度的不平衡性，出现了新老的分合是根据语言相异度还是历史文化性。《方言》屡次提到的与“南楚江湘”有别的“齐楚”语，是楚国后期扩张到淮北鲁南一带的东楚语。作为吴越和齐鲁过渡的东楚语，因地处淮水流域而与中原和齐鲁交通，汉魏西晋时不免受之同化。永嘉迁往江淮的北人多来自青徐，因此东晋北朝的东楚语与江淮语相近。徐淮的东楚语在隋唐进一步中原化、辽金汉语的阿尔泰化浪潮致使东楚语终与江淮语分道扬镳，然而江淮语的底层至今在徐淮方言中清晰可见。

湘赣浙南地名“洋、寮、圳”等，表明这一地区古属百越。客家祖先的南迁虽始于东晋，但民系意识的形成却在宋代。根据沙加尔的新说，唐初以前进入鄱赣的北方话和当地土语接触形成原始赣语。中晚唐的中原移民带来了新的北方话，和原始赣语接触后形成全浊声母消失的南赣语。随着赣中地区的北方移民经东南至闽西再转粤北，其南赣语和新住区土著语接触，在宋元明形成客家话。虽然客家话和南赣语的差别并不显著，但不存在同一民系意识，因此可划分为两种方言，而将北赣语独立成区。徽语是介于赣语、吴语和江淮语之间的丘陵盆地方言。至两晋之交出现中原汉人聚族迁人的浪潮，作为汉族民系的徽人在宋代已经形成。现代徽语的声母和韵母系统具异源性。根据江淮通泰话、赣客语和徽语全浊声母塞音、塞擦音今读送气清音的共同特征，可以推断其声母源于永嘉南迁汉语层；有些不见于浙北和苏南吴语，而见于浙南处衢片、东瓯片的韵母特点却可在

徽语里找到,表明徽语与浙南吴语包含着皖南——浙南土著的百越语底层。

无论是从民系还是方言特征,江淮语都不是现代或近古北方话的次方言,而是在永嘉以后就形成的一支汉语方言,官话是明代以南京话为标准的江淮语的通称。此区原是古吴楚语北境,永嘉和靖康南迁中原语言的两次深度覆盖,形成了保留中原语言不同历时特征的中部汉语方言。地处海隅的通泰,非兵家必争之地,语言年代较为古老。如果从淮河越长江、经浙南到闽地,上古都是东南百越的范围,那么中原南迁居民对东南百越语的影响则是由北向南依次减弱,古吴语北境的江淮全覆盖……古吴语南部的江南深融合……古越语的浙北浅融合……古瓯语的浙南深渗透……古闽语的闽地浅渗透。闽语并非从古吴越语中分化出来,而是中原汉语对闽地越语汉语化程度较低的结果。吴闽语的底层都是百越语,自然不排除存在内部差异。

错综复杂的因素决定了中原居民南迁的人口数和层垒性,中原首次移民某地的历史层次不等于后来形成的南方方言的语言年代。汉语南方方言的存古歧异性可用多种因素加以解释:1.就南迁汉语来说,中原人南下有先后时间差别,南下中原人原居地区有东西地域差别;2.就土著语言来说,与汉语互动的百越语有时间或地域的底层差别;3.就语言主体来说,各方言区中原汉人和当地土著人口有语主比例差别;4.就后世影响来说,再迁中原汉语对各方言影响的层次和程度具有差别;5.就方言演变来说,因周边语言环境不同而各方言具有流变差别。

有一个极具价值的现象必须注意。长江以南的史前南方印纹陶文化,其繁荣约当中原商周春秋,依据各自特征分为七个次文化区。除了江淮语插入宁镇,客家话分布在赣南、粤东和闽南,桂北土语群和平话分布在广西,史前南方土著印纹陶文化的

分区和汉语南方方言的分区具有对应性：1. 宁皖区（原始侗台族山越文化）——徽语（江淮语）；2. 太湖区（原始侗台族于越文化）——吴语；3. 赣鄱区（原始苗瑶族苗蛮文化）——赣语；4. 湖湘区（原始苗瑶族荆蛮文化）——湘语；5. 岭南区（原始侗台族南越文化）——粤语；6. 闽台区（原始侗台族闽越文化）；7. 粤东闽南区（原始侗台族骆越文化）——闽语[1]。南方印纹陶创造者的语言，就是分布在长江下中游以南的原始南岛语和南亚语。如果南方印纹陶文化中的不同亚文化相包含着不同的亚语言相，那么可以推定汉语南方方言的形成与当地土著的语言差异和人口密度相关，各文化区土著语的底层歧异性左右着后来形成的汉语南方方言的区域性差异。依据民族历史语言学和华夏汉语混成发生论，现代南方汉语方言的一些特征并非当时南迁中原汉语独有，而是中原汉语的前身华夏语的夷越语成分和原始侗台一苗瑶语所共有的同源性特征。

公元四世纪以后，北方游牧民族的南下引发一次次人口迁徙，长江中下游和越蛮土著的"交趾之会稽七八千里"的华南地带，逐步成为中原居民的新的家园。中国境内人口迁移的总体走向是：北方游牧民族由朔北大漠进踞河洛；中原汉人由河洛迁江淮和江南，或再南抵岭南；由此引发了百越迁移的连锁效应，导致了南方民族从江浙到闽粤，从长江流域到湘西川康，从珠江流域到云贵高原的迁徙走向。此起彼伏的人口迁徙留下了烙印，一方面，中国现代人口生物学树形图上突出的特征是——与一个汉族群体具有最密切关系的并非远地汉人，而是毗邻的其

[1] 杨锡璋《南方印纹陶和汉语南方方言》，《中原文物》1990年第2期。

他民族群体[1]。"汉族"不是一个以内部遗传为特征的民种概念，而是一个以共同文化渊源为心理基础的民族概念。另一方面，一部汉语方言形成史就是一部中原居民迁移与周边民族融合史，一部华夏汉语推移发展史。

与强调语言共时特征歧异而治丝愈棼的划分思路不同，汉语方言划分的基本原则应当是：1. 方言区域和移民轨迹、文化底层的一致性；2. 方言系统和汉语历史系统、周边民语系统的互证性；3. 语音系统、基本词汇和语法结构特征的综合性。全面而深入地研究方言，务必克服以普通话为主要参照系的定势，冲破汉语方言一元分化论，树立"混成发生·推移发展"的语言文化历史观，接受语言接触理论和社会语言学方法，注意发现汉语方言中的残留的非汉语底层，挖掘汉语在历史上接受的非汉因素。

秦汉以来，北方游牧民族呈周期性逐鹿中原，屡屡建立王朝统治中原乃至南北。从始皇结束封建联邦制统一中国，到清帝皇冠落地的2132年间，游牧民族统治中原和全国长达840年，这还不包括具有鲜卑血统的隋唐的326年。在北方民族入主之时中原士族大量南迁，在造成汉语主导语言的基础音系八、九百年一次历史性大变迁的同时，出现了北留汉语的阿尔泰化与南迁汉语的百越化，与之相应的是阿尔泰人的汉语换用与百越民的汉语渗透。毫无疑问，推移模式反映了周秦以后汉语与接触语言互动的基本态势，但只有在强大文明中心形成后，才有华夏汉语的推移性延伸。在强大文明中心的形成过程中，周秦以前的华夏语以混成趋势为主。如果不进行寻根推源式的探讨，中原汉语则成了无源之水。依据历史语言学的成果，与新石器时

① [美]王士元《语言的变异和语言的关系》，《汉语研究在海外》（石锋主编），北京语言学院出版社 1995 年

期考古文化系统、人种体质类型，与古史传说氏族系统等相互印证而提出华夏汉语混成发生论。在对桥本推移模式修正的基础上，建构"混成发生·推移演变"的理论模式。

东亚大陆语言发展的态势，可以设定远古存在与三大考古文化系统相应的三大太古语言。在民族接触和文化交流中，在三大考古文化或三大太古语言交汇的河洛地区，形成了一种与强大文明相伴随的混合型语言。这种混合型的华夏汉语，伴随着游牧文明与农耕文明的冲突交融及中原文明的辐射，像一个巨大的漩涡逐步地同化着所接触到的周边语言，并在这一历史过程中也改变着自己。周秦以前中原地区的语言发展主流是因交替换用而出现的混成，与华夏民族的融合形成、华夏文明的崛起相一致；秦汉以后汉语的发展因北方民族进踞中原、中原士族南迁而出现的推移，表现为中原北留汉语的阿尔泰化和南迁汉语的南方化及阿尔泰语的汉化和南方语言的汉化，与汉民族和汉文明的进一步延伸相一致。中华文明史就是一部农耕文明与游牧文明相互冲突并趋于融合的历史，华夏汉语史就是一部多种民族语混成和推移交融的历史。在接触语言绵延不绝的汉语化过程中，同化也就是异化。汉语史研究理论模式体现着一定的语史观，一定的语史观的深层是一定的文化史观，文化史观的转变会引起语史观及研究模式的嬗变。"混成发生·推移演变"模式的价值并不仅在汉语史领域，而且对中国文化史的重建也有不可低估的价值。

## 第二节　中国语的历史和历史的中国语

汉语具有广义与狭义的理解。狭义汉语指秦汉以后的汉人语言，以公元前汉王朝的建立为坐标，至今约有 2200 年。由周

秦上溯至黄帝进踞中原(约前 3000)其间约 2800 年的语言为华夏语。华夏语又可分为两大历史阶段:五帝时代(约前 3000～2200)和夏代(约前 2100～1700)为原始华夏语阶段;商代(约前 1700～1100)和周代(约前 1100～前 3 世纪)为商周华夏语阶段,因其语言表层不同又可分别称之为夷商语和姜周语。广义汉语包括华夏语和狭义汉语,或合称为华夏汉语。中华 5000 年文明史,即有相应的 5000 年华夏汉语史。

距今 7000 年左右,东亚大地上新石器时代三大考古文化系统大致形成。在前 5000 年到前 3000 年之间,对于这些远古先民所使用的语言,可以称之为中国太古语或东亚太古语。依据考古文化系统的命名,太古语可分别称之为青莲岗语、仰韶语与北方语;或依据见于中国古代文献的古代民族系统的命名,可分别称之为太古夷越—苗蛮语、太古氐羌语与太古胡狄语;或依据现代语言系属的命名,可分别称之为太古南岛—南亚语、太古藏缅语与太古阿尔泰语。所谓 7000 年中国语史,即指太古语、华夏语和汉语生生不息的历史。

根据文献语料,从周代开始,华夏汉语始终有一主导语言系统即官场、文学或商用的通行语。周秦为雅言,西汉为通语,东汉魏晋为洛生咏、北语,南朝为吴音,隋唐为汉音或秦音,宋元为宋音、中原正音或中原雅音,明清为官话、官语或官音。其中雅言、通语、秦音的基础音系是历史上的西安音,洛生咏、正音的基础音系是历史上的汴洛音,吴音、官话的基础音系是历史上的南京音。秦汉以后,汉语主导音系的变迁轨迹经历了西安音 → 汴洛音 → 南京音,又西安音→ 汴洛音 → 南京音的周而复始的变迁。而汴洛音是西安音在中原的后裔,南京音又是汴洛音在江淮的嬗变。至于清代末叶至 20 世纪逐渐上升为主导音系的北京音,则是明初迁往北京的江淮话在北方演变的结果。换而

言之，近现代北京音是南京音在燕蓟的后裔。

通过宏观探讨历史的华夏汉语溯源至中国太古语，以阐明7000年来中国语言之演变大势，可以进而提出东亚—南洋语言文化圈之假说，全景鸟瞰东亚—南洋诸语言(阿尔泰语·藏缅语·华夏汉语·苗瑶语·南亚语·侗台语·南岛语)的演变大势。

## 一、三大考古文化系统与三大太古语言区域

如果把36亿年生物史算做一年，每分钟大约等于7000年。在这个缩微时间内，12月31日猿类出现，12月31日晚上11点北京猿人开始用火[①]，晚上11点45分古人才完全直立步行，这以后前语言出现，而成熟的有声分节语言的历史只有5分钟左右。假若以上这些研究可信，东亚大陆人类的现代语言可能出现在3万年前，并由此引起旧石器时代革命而迎来新石器时代的曙光。

张光直曾提出，黄河流域前石器文化中有一个"尚未显出差别的汉藏南岛综合体"，从新石器时代起，开始向中国南部移民活动的结果分化出汉藏群和南岛群[②]。根据近30年来的考古发现，东亚新石器文化中心并非仅在黄河流域，处于东南的江淮文化与处于北方的松辽文化曾经高于黄河中游文化，反之，由于多种文化的冲突交汇才形成了后来居上的中原文化。这些不同的新石器文化是否源于同一旧石器文化母体，至少目前尚未找到证据。而竟成认为，汉语与阿尔泰语之间的底层关系可能反映了语系分化前这一地域的语言状况，因此有必要提出非谱系

① 参见[美]王士元《学术报告》，《语言学论丛》第11辑，商务印书馆1983年。

② [美]张光直《华南史前民族文化史提纲》，台湾《民族学研究所集刊》1959年第7期。

性的历史语言学概念[1]。随着文明的传播与交融，由非谱系语网形成太古语区，然而推断太古语区的形成时间必须找到一个支点。

石兴邦认为："考古学上的文化体系与民族学上的涵义是相同的。它是人类社会生活各个方面系统化的综合体。在这个综合体中包括一系列具有特征的文化因素，像社会组织制度、民族的思想行为和方式、生活习俗和风尚、工艺技术特点和风格，以至宗教信仰、道德规范等社会行为模式。这些因素，互相融铸而形成一种不可分割的有机联系的整体，成为所属社会特有的文化形态。"[2] 然而，无论是作为文化的语言，还是作为工具的语言，在考古文化系统综合体中都是不可缺少且具有显著特征的基本要素之一。因此，当假定农业社会之前是一个尚未显出差别的网状文化体，而定居农业的出现则是促使网状文化体分化的历史条件，并由此进一步形成考古文化系统和太古民族系统，也就为太古语区的形成找到了可以尝试探测的支点。

考古文化研究表明，大约距今 7000 年即公元前 5000 年左右，中国境内的新石器时代三大考古文化体系大致形成，并由此奠定了中国境内历史文化发展的结构和态势。处于东南沿海及湖泽平原的是水稻培植为主的青莲岗水耕文化系统；处于西北黄土河谷地带的是以旱粟种植为主的仰韶旱耕文化系统；处于河套大漠草原地带的是以游牧为主的北方细石器文化系统[3]。三大考古文化系统的创造者都是蒙古利亚人种，但体质特征有别。东南文化的创造者是南亚型，即太古夷越人；西北文化的创

① 龚成《汉语史研究的新思路》，《现代语言学》1992 年总第 21 期；《古代汉语无音和谐现象》，《探索与争鸣》1986 年增刊。

②③ 石兴邦《中国新石器时代考古文化体系及相关问题》。

造者新石器时代是东亚型与南亚型，青铜时代是东亚型，即太古氐羌人；北方文化的创造者早期为北亚型，晚期有北亚型和东亚型混合的趋势，即太古胡狄人[1]。可以推定，与三大考古文化系统的形成相平行，太古青莲岗语区、太古仰韶语区与太古北方语区等三大太古语区也在同时形成。

与三大考古文化系统相印证的是中国文献古史中的三大传说氏族系统。20世纪二、三十年代，以顾颉刚为首的疑古派滥用默证法，把大禹考成一条虫，把古史传说斥之伪古史，5000年文明史一笔砍去2000年。上古远史口耳相传，辅以原始实物（结绳、刻划、图画）记录，至文字发明始书之典册。尽管在口传与笔削时会有神化、夸张、讹变、附会、删并、嫁接等，但决不可视为子虚乌有而全盘否定。随着三大考古文化系统的发现，三大传说氏族系统得到印证。兴于江淮而发展至黄河下游的伏羲氏，为夷越苗蛮之始祖。兴于渭水而发展至黄河中游的神农氏，为氐羌之始祖。兴于北方草原而后南下中原的黄帝，为胡狄或阿尔泰之始祖[2]。游牧民族被称为"马背上的民族"而最早发明马车，传说轩辕氏黄帝造车。伏羲氏称太皓、神农氏称炎帝、轩辕氏称黄帝，"皓、炎、黄"皆指阳光，反映了东亚先民的太阳神崇拜。前3000年左右存在的三大传说原始氏族的语言，可分别称之为原始夷越语、原始氐羌语和原始胡狄语。

石兴邦认为："从考古文化的观点说，在一个文化体系中，还包含若干个不同的文化共同体。这些文化共同体是由该氏族部落聚居的地区、时间和从事经济生活性质的异同而形成特有的

---

① 格勒《中华大地上的三大考古文化系统和民族系统》。

② 郭沫若主编《中国史稿》(1)，人民出版社1979年。

文化相。"[1] 因此,在每一远古语言区域内也包含着若干个不同的语言共同体,即由聚居地区、时间和经济文化因素而形成的若干特有语言相,即通常所说的语群和语种。原始夷越语系可能包括九夷语、百越语、南蛮语、百濮语等;原始氐羌语系可能包括氐语、羌语等;原始胡狄语系可能包括西戎语、北狄语等。这些有差别的原始语又成为后来语言分化的源点。现代阿尔泰语、现代南岛语是粘着型,现代南亚含有粘着成分,所谓语系不明的朝鲜语与日本语也是粘着型。至于现代藏缅语,藏语的句法方面更像阿尔泰语,藏语在历史上经历过粘着型→屈折型的变化[2]。其中有些语言(如嘉戎语、僜语、景颇语)的附加成份十分丰富,可以认为原始藏缅语的古老结构是粘着型。受这些粘着型语言包围的现代汉语却为孤立型,盖是语言多次混合的结果,原始华夏语应为粘着型。侗台语、苗瑶语中也有不少粘附形式,其孤立型倾向的成因可能是受汉语影响所致[3],其古老类型为粘着结构。可以推断,形成于前5000年左右的三大太古语与发展至前3000年左右的三大原始语言,其语言类型可能都是原始粘着型。

## 二、华夏语的混成发生与类型的动态转变

考古研究表明,青莲岗文化的一支溯黄河而上,仰韶文化的一支沿黄河而下。经过1000多年的频繁接触,大约前4000年左右在河南地区形成了一个新的共同体——庙底沟氏族文化。

---

① 石兴邦《中国新石器时代考古文化体系及相关问题》。

② 胡坦《国外藏语语法研究述评》,《国外语言学》1993年第2期。

③ [苏俄]E. B. Иванова《论台语及其在发生学分类法中的地位》(1970,许浩福译),《汉藏语系语言学论文选译》,中国社科院民族所语言室1980年印行。

这个文化共同体包含有两大文化系统的一些共同特征，但以东方文化占优势，即与青莲岗文化较接近，而仰韶文化少些[①]。庙底沟氏族文化形成以后，就以旺盛的生命力有力地融合着附近的部落文化，逐步占有了原仰韶氏族部落文化的区域，使中原地区成为文化融合的漩涡，将新的文化因素吸收进来又洪溢出去，致使融合性的文化因素向边缘区域漫延，由此出现了最早的混合型中原文化的原始共同体。

三大考古文化系统的逐步混合过程，积淀在古史传说之中。文明的传播、洪溢与混合，在史前往往通过战争方式为媒介。古史传说中有三次较大的氏族战争。第一次是发生在前3000年之际的东南蚩尤与西北共工之战，暗示了东南夷越文明与西北氐羌文明的早期接触，或许是庙底沟文化共同体形成在传说中的折光。第二次是北方黄帝与蚩尤之战。共工部落战败后，炎帝北上求援，结成炎黄联盟以对付蚩尤的传说反映了西北农耕文明与北方游牧文明的早期接触。黄帝擒杀蚩尤于冀州之野，并令少皞统领夷人各部，反映了北方游牧文明与东南农耕文明的接触。第三次是炎帝与黄帝之战，炎帝欲争盟主地位，黄帝与之战于坂泉，"三战，然后得其志"。胜利了的黄帝乘势向南发展，"五十二战而天下咸服"。黄帝进踞中原、直至江汉，为三大原始文化系统的加速融合创造了历史契机，为华夏民族的形成和中原文明的发展奠定了基础，由此黄帝被奉为"人文初祖"，司马迁《史记》从"黄帝"落笔以溯华夏文明之源。

人类文明的演进史表明，封闭的、单纯的文化往往是迟钝的、落后的文化，而开放的、混合的文明往往是进取的、发达的文明。巴比伦文明承袭光大了苏美尔文明与阿卡德文明，而古希

---

① 石兴邦《中国新石器时代考古文化体系及相关问题》。

腊文明更是建立在米诺斯、迈锡尼文明的废墟之上。在中华文明史研究中，传统的单源优势论长期占据着主导地位。近年来，有人提出所谓"一体多元论"，然不经多元聚合何来一体？由数种文明冲突而后整合混成的文明往往是强大的文明，这种强大文明中心形成以后，又以洪溢方式持续扩展，从而形成主导文明和区域文明互动发展的格局。三大考古文化系统的发现和中原原始文化共同体的发现，找到了华夏文明的"原生老根"及其生成发展的历史氛围。既然华夏民族奉黄帝为远古先祖，那么，原始华夏语的源头亦当追溯至5000年前，即以黄帝进居中原为起点。自黄帝以后，经颛顼（夷越系）、帝喾（夷越系）、尧（胡狄系）、舜（夷越系）的交替统治，华夏民族与华夏文明在中原地区形成，原始夷越语、氐羌语、胡狄语也逐步混成原始华夏语。历史传说中的氏族首领谱系的混杂性和氏族首领居住地的变迁性，正反映了原始华夏文化的互渗混成性。

华夏语经夏商周三代而成熟。夏代无文献传世，夏语难以详论。《说文》："夏，中国之人也"，中原居民称"夏"始于夏代，因而夏王朝对华夏族的发展起着承上启下的作用。夏禹又称戎禹，《诗经·长发》记载："禹敷下土方"。据甲骨文推测土方在商之西北的河套一带。夏人由西北进居中原，可推知其为胡狄系。《史记·匈奴传》："匈奴，其先夏后氏之苗裔也，曰淳维。唐、虞以上，有山戎、猃狁、荤粥居于北蛮，随畜牧而转称。"匈奴又译为"胡"。上古音"胡" * ɣa、"夏" * ɣea 十分接近[1]，其差别在等呼，若根据古音无等呼则同音，由此可见夏语与胡语关系密切。又《史记·六国年表》记载"禹兴於西羌"，似乎暗示以胡狄为主的夏禹，在进入中原前已与氐羌混血。进入中原后难免又与夷越交

1 史有为《异文化的使者——外来词》，吉林教育出版社1990年。

融，因此戎夏语可能是混有氐羌和夷越的原始胡狄语。

商人又称“殷”，“殷”与“夷”、“人”古音相同。《说文》：“夷，东方之人也”，《诗经·玄鸟》：“天命玄鸟，降而生商”，鸟崇拜是原始夷越文明的特色之一。商人以鸟为祖先之源，已为甲骨文所证实，商人无疑是以鸟为图腾的一支东夷。但《史记·殷本记》记载“殷契，母曰简狄，有娀氏之女”，《诗经·长发》记载“有娀方将，帝立子生商”，表明了商人承袭了戎夏文明因素，商人祖先是以东夷为主而融进戎夏的混血。因此夷商语可能是掺入胡狄、氐羌的原始夷越语。商王所拥有的是邦联政权，见于甲骨文的邦国有 90 多个[①]。商之北有以居方为首的胡狄系部落，商之西有以周方为首的氐羌系部落，商之东南亦当九夷、徐夷等部落。胡厚宣认为：“商而称中商者，当即后世中国称谓之起源也。”[②] 商人的语言中，人名多祖甲、祖乙，地名多丘商、丘雷，病名为疾口、疾耳，此为南方语言特征的 NA 顺行结构。甲骨文语法中，动宾补的语序竟然 6 种可能有的语序都存在[③]。对于不以词序为语法手段的解释只有一种可能，即这是一种具有形态手段的语言，夷商语可能还是粘着结构，只是这种结构的粘着方式被文字的表词性质掩盖了。

来自西方的周人，其祖母为姜原。“姜”、“羌”同源，《说文》：“羌，西戎牧羊人也”。《国语·周语》：“昔我王先世后稷，以服事虞、夏，及夏之衰也，弃稷弗务。我先王不用失其官，而自窜于戎狄之间”，周人又承袭黄帝族姓之“姬”，盖以氐羌为主而与胡狄

① 赵诚《商代社会性质探索》，《古代文字音韵论文集》，中华书局 1991 年。

② 胡厚宣《论五方观念及中国称谓之起源》，《甲骨学商史论丛》，台北大通出版公司。

③ 余志鸿《论古汉语补语的移位》，《语言研究》1984 年第 6 期。

混血。其同宗称"诸夏",其姜周语名"雅(夏)言"。俞敏以为,藏语与周代姜姓独有的方言更像,汉藏两族人和话同源[①]。然而周人既承袭了中原殷商文化且借用了殷商文字,因而其雅言必与殷商语接触交融。周人在以河洛为中心的广袤疆域建立了封建联邦制,通过分封诸夏、移民洛邑、礼乐教化、编撰典籍以传播雅言而使中原语言氐羌化。土族语言 → 官场语言 → 礼教语言 → 文学语言,这就是上古雅言形成和扩展的轨迹。传说周公撰《尔雅》(迩夏),其旨趣在促使方(邦)言向雅言靠近。尽管春秋战国礼崩乐坏,但雅言作为华夏语言文化的象征已成共识,"五经"《易》(巫术哲理)、《书》(传说历史)、《诗》(民风文学)、《礼》(宗法伦理)、《乐》(教化音乐)皆撰成于周。这些氐羌化的书面语言,作为华夏文献文化的源头,对后世文化发展之影响非常深远。

关于三代华夏语的异质混成性,白保罗曾提出,汉藏语的成份只构成汉语的表层,而底层另有不同来源,周朝人所操的某种汉藏语融合或渗入到商朝人所操的非汉藏语言之中[②]。蒲立本(E. G. Pulleyblank)曾主张夏是古汉语族,商与南亚语系关系密切,而周之先世是藏缅语族[③],皆未见其详细论证。在历史比较语言学研究中,不管是汉—藏同源说、汉—南(南岛)同源说[④],还是

① 俞敏《汉藏两族人和话同源探索》,《北京师范大学学报》1980年第1期。

② [美]P.K. Benedict: *Sino-Tibetan: A Conspectus*, Cambridge University Press, 1972。中译本《汉藏语言学概论》(乐赛月、罗美珍译)。

③ 见邓晓华《人类文化语言学》,厦门大学出版社1993年。

④ [法]L. Sagart: *Chinese and Austronesian are Genetically Related*, 23rd International Conference on Sino-Tibetan languages and Linguistics, 1990。邢公畹《关于汉语南岛语的发生学关系问题》,《民族语文》1991年第3期。

汉—阿(阿尔泰)同源说[1],都是依据各自选择的语料及人种学、文化史方面的材料,来论证汉语与某一毗邻语的发生学关系。华夏民族的形成是史前氏族融合的结果,与之同步形成发展的华夏语不可能不具有与被融合语相一致的成分。与其说汉语仅仅与某一毗邻语同源,不如说在历史发展的长河中,后来被称为“华夏汉语”的这种语言与毗邻三大语言都先后发生过渗透、换用与混合关系,华夏语是在多种原始文化撞击而整合为中原文明的历史互动中产生的历史混合语。

关于周秦语言的研究,近年来已经取得许多明显进展。一些问题已渐趋明朗:1. 有复辅音;2. 无音位性声调(可能有伴随性声调);3. 闭音节较多或辅音尾较多;4. 音节结构通常是CCVC型。据沙加尔(1990)的研究,原始南岛语(相当于原始夷越语)的语词是复音节的,它的不少后代的音节缩短了,上古汉语是典型的单音节语词结构,汉语的“字”与原始南岛语对应的经常是语词的末一音节,如:“白” * burak(原始南岛)< * brak(上古汉语) < * bak(中古汉语);又如:“刮” * kurat(原始南岛) < * kwrat(上古汉语)< * kwat(中古汉语)。由此揭示了华夏汉语语词音节结构由CVCVC到CCVC,再到CVC的演变过程,但是单音节化过程在各方言中的发展是不平衡的。沙加尔引白保罗之说,认为南岛语的bulut“纤维;软毛”可和汉语的“笔”对应,郑张尚芳认为印尼语balut“涂抹”与汉语“笔”对应

---

① 参见岑仲勉《揭出中华民族与突厥族之密切关系》,《东方杂志》1945年41:5;《楚辞注要翻案的有几十条——楚辞中的古突厥语》,中山大学学报1961年第2期。赵相如《维吾尔语的音节结构和借词拼写法的关系》,《民族语文》1984年第4期;《汉语“氏”和维吾尔语“tag”的关系初探——兼论汉语阴声韵的辅音尾》,《延边大学学报》1984年第4期;《维吾尔语中的古代汉语借词——从语言上看历史上维汉人民的密切关系》,《语言与翻译》1986年第2期。

(邢公畹 1991)。《说文》"聿,所以书也。楚谓之聿,吴谓之不律,燕谓之弗","笔,秦谓之笔"。当时秦语 * pljiət、楚语 * ruət、燕语 * pjət,而吴语 * pjəu-ljuət。前三者已单音节化,而吴语仍为双音节。东晋郭璞《尔雅注》:"蜀人呼笔为不律也",蜀语其时仍为双音节。

蒲立本等人推测古代汉语曾经有过丰富的形态,试图建立汉语历史形态学。同族词的研究表明,古代汉语构词多用语音交替[①]。而潘悟云提出建立汉语历史形态学有"两个窗口",一是对异读、假借、谐声的系统研究,一是亲属语的历史比较[②]。然而,必须设定华夏汉语结构类型的动态演变,为历史形态构拟提供依据。我的初步预设是华夏汉语的语言类型经历过粘着型→不完全屈折型→孤立型的动态演变。商代华夏语还是粘着型,粘着性词缀可能有辅音性和音节性两种,周代华夏语发生单音节化且屈折方法产生,随着形态的纷纷脱落,华夏语向孤立型方向演进。复辅音蜕变为单辅音,韵尾辅音失落,音位性声调出现同时复音节合成词增多。在华夏汉语类型孤立化进程中,一些音节性词缀与实词蜕变为虚字。这一设想与严学窘大致相合[③]。在谱系树模式占统治的时代,不恰当地强调语言的一源分化,而无视语言的多源聚合。混合语既是语言混成的结果,又是文化交融的表徵。语言学上有一条规律,语言成份越混杂,形态丢失越多,克里奥耳语的形态都大大少于源语[④]。多向性、层叠性的混合促使华夏汉语语言类型由粘着,经不完全屈折演变

① 王力《同源字典》,商务印书馆 1982 年。

② 潘悟云《第三代汉语音韵学》,《现代语言学》,延边大学出版社 1990 年。

③ 严学窘《原始汉语的研究方向》,《王力先生纪念论文集》,商务印书馆 1990 年。

④ 刘丹青《文化研究中的"语言导人"》,《东南文化》1991 年第 1 期。

为孤立语类型。

## 三、南耕北牧生态环境中的汉语嬗变轨迹

公元前后，欧亚大陆以联绵于中部的诸山脉（兴安岭—阴山—祁连山—阿尔泰山—天山—昆仑山—兴都库什山—札格罗斯山—高加索山—陶鲁斯山—喀尔巴阡山）为界，形成了南耕北牧的文明格局。农耕世界以农业经济和城市文明为主，富裕而保守；游牧世界以游牧经济为主、文明程度较低，但机动性强，富有开拓精神。两个世界一经形成，冲突频繁发生。总的趋势是游牧民族处于攻势，而农耕民族处于守势。当某个游牧民族在强大的头领的率领下统一起来时，一次大规模的民族战争就来临了。进入农耕世界的征服者都不得不适应当地经济文化方式，而为被征服者所同化，并且大部分还不得不使用被征服者的语言。当农耕世界恢复平静而民族融合渐趋完成时，战争的乌云又在山那边重新聚集，一场新的民族战争又在孕育之中。游牧民族由于渴望享受文明成果而发动的进攻，成为世界上一个循环往复的问题。征服、迁徙、融合，再征服、再迁徙、再融合……成为古代民族史的主旋律[1]。与之相应，毗邻语言之间的接触、融合、再接触、再融合……亦成为古代语言史的主旋律。在这一语言融合交响乐中，对于进踞农耕世界的征服者来说是语言的换用，对于被征服者的农耕世界的语言来说则是语主的大规模性嬗变，从而使语言变化的矢量表现为时间·空间·主体的三维互动。

在东亚大陆，农耕文明与游牧文明的冲突融合在史前已经

① 商友仁《古代欧亚大陆民族迁徙和民族融合鸟瞰》，《北方论丛》1988年第4期。

发生,并由此触发了华夏中原文明的产生并形成了混成性的华夏语。秦汉以后,这一南耕北牧的冲突交融更为广阔、持久和深入。在北方先后继起的游牧民族有匈奴、鲜卑、柔然、突厥、契丹、党项、女真、蒙古和满洲,尽管中原汉人以征战、长城、联姻、册封、内迁等方式阻挡游牧民族南下,但是秦汉以后,北方民族入主中原或统治中国全境还是长达840年之久,这还未包括含有鲜卑血统的隋唐王室的几百年统治。以永嘉之乱、安史之乱和靖康之难为高潮,北方游牧民族的进踞中原引发了中原汉人的一次次大规模南迁,从而导致了"北留中原汉语"的阿尔泰化或进踞中原的阿尔泰人的语言换用,与"南迁中原汉语"的夷越化或接受同化的夷越土著的语言换用。汉文化日益延伸,汉人口日益壮大,而汉语在持续同化周边语言的过程中也异化着自己。

在这一绵延2200余年的历史长河中,汉语主导语言的基础音系也实现了一次又一次的变更。秦都咸阳,"书同文"以秦文即上古秦音为正。西汉都长安,其主导语言的基础音系承袭周秦雅言而演变为通语,同时东南地区的楚文化复苏导致汉语的嬗变。东汉都洛阳,曹魏、西晋相沿都于河洛,以洛阳语音为主导语言的基础音系历时300年,形成了魏晋文学标准音洛生咏。周公建洛邑为"成周",平王东迁于此,洛阳音可溯源于周代镐京王畿之音,但东汉魏晋洛阳音与西汉长安音当有地域和时间的差别。永嘉南渡,洛中人士迁往江东,建康风行洛生咏,同时中原人士亦仿习吴语。经过宋齐梁陈,源于中原而兴于江东的江南文化日益成熟,源于洛阳而在江南发展的金陵音,成为新的主导语言基础音系,日本奈良时代到7世纪输入的六朝汉语字音称为吴音,唐人亦如此称呼。两汉魏晋经东晋到南北朝,汉语主导语言的基础音系展现出长安→洛阳→金陵的横折型大变迁,

历时800年之久(前206～后600)而实现了汉语从上古至中古的嬗变。

隋唐建都长安,《切韵》虽折衷于金陵、洛下,但大唐的政治文化以长安为中心,主导语言的基础音系难免西移,这从李涪斥《切韵》吴音乖谬略见一斑。天宝年间(742—755),元庭坚更一反"时俗共重以为典则"的《切韵》,直以秦音撰成《韵英》。至慧琳《音义》(807),音切一以《韵英》为则,斥陆韵为"吴楚之音"。日本遣唐学生和僧人将长安方言作为正雅之音带回日本,时称为"汉音"并逐渐排斥六朝旧音,有"诵两京之音韵,改三吴之讹响"之说[①]。五代建都汴京,北宋亦都于此而名东京,主导语言基础音系的东移形成了宋元明所称的"正音"或"中原雅音"。靖康之难导致中原居民南迁,除随宋室往江浙临安,更多的中原居民留居在以南京为中心的江淮一带。日本镰仓时代由江南输入的汉字音称之为"宋音",即指由中原迁往江南的北宋正音。迁往临安的宋音,虽保留北方话的一些特点而终究为周围吴语同化。迁往江淮之间的中原正音,因江淮语在六朝时已被中原汉语同化且处于吴语北部边缘,所以正音相对得以保存[②]。据元戚光《集庆路续志》:"金陵山川浑深,土壤平厚。在宋建炎中城境为墟,来居者多汴洛力能还迁巨族士家,视东晋为又一变。岁时礼节饮食,市井负衔讴歌,尚作京城故事",当时建康多为汴洛流民[③]。元代所撰《蒙古字韵》,日本汉学家服部四郎认为记录的是迁徙到临安的汴京方言,而中野美代子《八思巴字与蒙古字

① 尉迟治平《"上声厉而举"解》,《音韵学研究》第3辑,中华书局1984年。

② 张卫东《论客家文化的直接源头》,第2届全国语言与文化研究会论文,1991年。

③ 鲍明炜《南京方言历史演变初探》,《语言研究集刊》,江苏教育出版社1986年

韵语音研究》认为这一音系中渗有吴语方言成分[1]。元亡明兴,留居汴洛的中原之音,历辽金元近300年统治已非北宋之中原正音,难免"北杂夷虏";而迁往临安的亦已"南浸吴越",唯有江淮之间的中原之音相对纯正。明初所修《洪武正韵》"壹以中原雅音为正",此"中原"已非地理概念,而是"中原传统文化"概念。所谓"中原雅音"即指保留在以南京音为代表的江淮话中的中原旧音,并以此成为明清官话的基础音系。日本江户时代传入的汉语字音为南京官话和杭州官话语音,而称之为"唐音"。隋唐北宋经南宋到元明,汉语主导语言的基础音系又展现出长安→洛阳→金陵的横折型大变迁,历时900年之久(600—1468)而实现了汉语从中古至近古的嬗变。

秦汉以后,汉语主导语言基础音系的巨大变迁,出现了由西至东再南移的横折型循环趋势。这种近乎八、九百一次的音系演变大循环与游牧民族入主中原密切相关,而游牧民族的进攻征战与牧草丰歉、畜牧兴衰有关,而牧草丰歉、畜牧兴衰与草原生态气候的周期变化有关。永嘉之乱,一方面洛音流播于江淮,另一方面胡人"断诸北语,一从正音",汉语"南染吴越,北杂夷虏",为中原之音一大巨变。隋初长安论韵,陆氏"捃选精切,除削疏缓",意在维护金洛之正音。安史之乱、黄巢之乱,造成大批中原居民南迁,五代沙陀人建都汴洛,为中原之音又一巨变。北宋颁行《广韵》等,其旨"设教崇文,悬科取士",兢兢不失其坠。靖康之难,中原之音四分五裂,有迁往江浙后为吴语同化之临安京音,有携往辽金而为近古燕代方言之底层者,亦有留于中原未

---

① Miyoko Nakano: *A Phonological Study in the Phags-Pa Script and the MengKu Tzuyun*, Faculty of Asian Studies in Association with Australian National University Press, Canberra,1971.

迁后演变为近古汴洛方言者，更有迁往江淮而后通行为明清官话者。此为中原之音又一巨变。明初所修《正韵》"复恐拘于方言"，因此"壹以中原雅音为定"。这种民族与民族、语言和语言反复融合的结果，则是淮河以北的北方汉语迅速变化并日趋统一，长江以南的南方汉语因山川阻隔而多存古语，中原士人大量迁居的江淮之间则保留了相对纯正的中原正音。由此形成了桥本万太郎所说的时间上由今及古、地域上由北而南的语言推移演变的连续体状态。

## 四、明清官话的市民社会内涵及其基础方言

"官话"之名通行于明代，或称"官语"。朝鲜《李朝实录·成宗四十一年)九月》载："头目葛贵见《直解小学》曰：反译甚好，而间有古语，不合时用，且不是官话，无人认听。"① 《直解小学》的作者是偰长寿，此书为李朝初期广泛使用的汉语教科书。这是目前所知"官话"一词的最早出处。1483 年"官话"已见于域外文献，由此推测明初已通行"官话"一词。

山东临清人谢榛(1495—1575)的《四溟诗话·卷三》、江苏华亭人何良俊(1506—1573)的《四友斋丛说·史十一》、江阴人徐充的《暖姝由笔·方言》、江宁人顾起元(1565—1628)的《客座赘语》、浙江乌程人凌濛初(1584—1644)的《二刻拍案惊奇·卷三十七》皆提到"官话"的使用和特点。而江西新建人张位(1582 前后)的描写较为明确，《问奇集·各地乡音》曰："大约江北人声多作平声，常有音无字，不能具载。江南多患齿音不清，然此亦官话中乡音耳。若其各处土语，更未易通也。"江淮间至今尚有人

① ［日］太田辰夫《关于汉儿言语——试论白话发展史》，《汉语史通考》(江蓝生、白维国译)，重庆出版社 1991 年。

声，江北当指淮北以远。官话是指含有入声的江南（明代江南行中书省，辖境相当今沪、苏、皖等地）话，即以南京音为标准的明代江淮话。所谓"齿音不清"，盖指正齿、齿头二音不别，但这是官话中的乡音。

明末西洋传教士来华所习之"官话"，范礼安（Valignano）称之为 common language（共同语）、the official language mandarins and of the court（官吏和朝廷的公务语言）；罗明坚（Ruggieri）称之为 the court language that they call mandarin（中国人称之为官场话的朝廷语）；利玛窦（Ricci）称之为 a universal language（通用语），并说 we can call it the language of the forensic court（我们可以叫它为衙门语言）；后来，他知道中国对此种通行语早有专名，在 Storia 中将"官话"拼作 Cuon-hoa，后又拼作 Quon-hoa。[①] 杨福绵认为金尼阁把利玛窦 Storia 中的原文"地道的中国语"改成"纯粹的南京话"，可证南京话就是当时官话的基础方言，可见明代两京和鲁赣吴地区及域外人所习之官话即南京话。

官话之"官"，在明代并不仅指官场、官司，或西洋传教士所说 mandarin 即官吏之"官"。"官"之初义治众之所，引申为治众之人，又君主、尊长之敬称。南朝时引伸为王族及士大夫年轻子弟之称。唐时惟有居官者方称"官人"，宋乃不然，对男子皆可敬称为官人。奴婢称主人为官人，妇人称丈夫为官人，故"当时殆无不官人者矣"，亦可单称为"官"。宋元以后，江南商业经济蓬勃发展，市民社会形成，商贾称朝奉、员外于乡里，普通百姓被称

① ［美］Paul Fu-mie Yang: *The Portuguese-Chinese Dictionary of Matteo Ricci: A Historical and Linguistic Introduction*, The Second International Conference Sinology Taipei, Taiwan, 1986；曾晓渝《试论〈西儒耳目资〉的语音基础及明代官话的标准音》，《西南师范大学学报》1991 年第 1 期。

为看官、客官，“官”的称呼已经市民化，犹如今日之“老板”。明清通用语称之为“官话”，其中包含着市民社会的丰富文化内涵，官话的出现及通行与宋元以来江南商业经济的繁荣密切相关。官话不仅用于官场、官司，不仅用于诵佛、吟诗，更广泛的是用于商贾贸易。宋代的话本、元代的戏曲、明代的评话小说，这些与百姓生活忧乐与共的市民文艺，对官话的兴起、流行、传播起着政府所不能起的巨大作用。因此明代的官话与前代的雅言、通语、洛生咏、正音的性质不完全相同，它不再局限于朝廷用语、公文辞章与文人吟诗作赋，而是有着广泛的市民社会基础、为社会各界所使用的通行语。

由于靖康之难汴洛居民多迁江淮，由于宋元以来江南商业经济的蓬勃发展，由于元末农民起义多为江淮居民且朱元璋祖籍金陵之句容，由于江南（行省）市民文化日益发达，因此，明清官话的基础方言势必是以南京话为代表的江淮方言。在中国历史上，入主中原的北方游牧民族多仰慕汉文化并主动同化。唯有蒙古元对汉文化尽力排斥，又人分四等而“南人”最下，并企图以八思巴字拼写元朝境内一切语言而取代汉字。周德清为捍卫汉字读音而撰《中原正音》，在书中集中刊布了《看岳王传》、《误国贼秦桧》等曲文，又虚拟沈约以所生地吴兴之音制韵一事而指桑骂槐：“宁忍弱其本朝，而以敌国中原之音为正耶？”[1] 周氏长期淹留江淮，坚守中原正音之民族立场，岂会以去都没有去过的“元大都”之音为标准？脱离历史背景和文化心态，孤立地研究《中原》音系，强行将《中原》附会元代大都音，势必致使近代汉语语音史被扭曲。明初《洪武正韵》（1375）所记录的是以当时南京

① 赵诚《周德清和中原音韵》，《古代文字音韵论文集》，中华书局1991年。

音为主体的江淮方言音系或南方官话音系①。朝鲜李朝撰修《洪武正韵译训》(1455)作为学习汉语的权威书籍,而将燕京北音以"俗音"附记于正音之下。并曾多次派人往燕京向汉儒求教,"以尽正俗异同之变"②。远藤光晓根据《翻译老乞大·朴通事》里的汉语声调的研究,认为属于下江官话的南京话在当时最有可能占有标准音的地位,北京一带当时通行来自南京的官话③。鲁国尧根据利玛窦《中国札记》的记载,推定明代官话的基础方言为当时南京音④。此后又有多人论证利玛窦、金尼阁所记官话音系为当时南京音([美]杨福绵、曾晓渝、张卫东1991,[韩]金薰镐1994),并试图从汉语演变史的角度阐述明代官话是以南京话为代表的江淮官话(薛凤生1992,李葆嘉1992,邓兴锋1992)。

除了明代官话基础方言的研究之外,清代官话基础音系仍是南京话的证据也已发现。18世纪中叶,莱阳客商白世云雇了常熟的船,从山东将豆子贩往江南,半洋忽遇风暴而漂至琉球,为琉球人学汉语编写了一本《白姓官话》(序于乾隆五年,1753)。濑户口律子的研究表明,《白姓官话》是以南京话为基础的语言⑤。日本天理大学所藏《白姓官话》的一些字下注有同音字,如:棒半、滩汤、参生、民明、圣甚。对于这种-n/-ŋ不分的现象,

---

① [日]满田新造《中国音韵史论考》,东京武藏野书院1964年。

② [韩]蔡瑛纯《译语类解所见中国音系之研究》,台湾中国文化大学硕士论文,1978年。

③ [日]远藤光晓《翻译老乞大、朴通事里的汉语声调》,《语言学论丛》第13辑,1984年。

④ 鲁国尧《明代官话及其基础方言问题》,《南京大学学报》1985年第4期。

⑤ [日]濑户口律子《从声调上推测〈琉球白姓官话〉的方言性质》,大东文化大学《语言教育研究论丛》1991年第8号。

濑户口律子依据《江苏省和上海市方言概况》中所记南京方言-n／-ŋ之分而不解。其实，南京方言老派至今“参—生；民—明；滩—汤”尚不分，南京周围以及江淮方言都可以发现这一特点的存在。由此证明，清代中叶的官话基础音系仍是南京话，不仅山东商人以之为通行语，就是域外所习也以此为标准。在江户时代（1603—1867），由黄檗宗的僧人、长崎的通事及民间商人，传入日本的汉字音称之为唐音，这种汉字音的语言文化背景就是明代至清代中期（14～18世纪）的南京官话，《白姓官话》可与之互为验证。清代官话亦以南京话为基础方言的明确佐证是，日本明治九年（1876）才确定，把以往学习对象“南京语”改为“北京语”。后以“北京官话”为代表性的称呼，有时也叫“清国官话”、“支那官话”，单纯用“官话”指“北京官话”是日俄战争前后（明治三十七年，1904）的事[①]。1878—1982年，第一位受聘于美国哈佛大学的徽州籍举人戈鲲化（1835—1882）所教汉语即为南京话。当时美国《每日图文》（1879年11月3日）介绍：“官话是全中国普遍使用的官方语言，是中国文学的载体，也是贸易界使用的语言。”[②]

至清末，京音地位有所上升，但尚未得到士宦界的公认。赵元任儿时随祖父居河北，学京音被其兄斥之为“干吗学老妈子说话的声音”，傅斯林家中亦“不屑说北京话”[③]。清末切字运动，卢戆章主张“以南京话为通行之正字，为各省之正音，则十九省

① ［日］六角恒广《日本中国语教育史研究》（王顺洪译），北京语言学院出版社1992年；［日］牛岛德次《日本汉语语法研究史》（甄岳刚编译），北京语言学院出版社1993年。

② ［美］崔颂人《美国汉语教学的先驱——戈鲲化》，《世界汉语教学》1994年第3期

③ 赵元任《我的语言自传》，台湾《史语集刊》43：3，1971年。

……文话皆相通。"但北方人主张"京音"为民国"官话"的标准,多方争执不下,遂有1913年"读音统一会"每省一票、服从多数、逐字审音之公决。老国音不伦不类,京音派渐占上风。由于北方游牧民族宋元后多次入主中原,打破了中原文明的原有一元格局,造成了经济、文化南移而政治中心北上的二元新格局。元明清三朝定都北京凡700年之久,为"京音"有可能确定为新的标准音奠定了历史基础。适逢北洋政府的首都在北京,国语统一筹备会1924年才决定修改老国音,以北京音为标准。又经过近几十年来以行政方式推广国语或普通话的运动,北京音才逐步取得通行的国音地位。然而,并不能因为20世纪以行政手段颁布北京语音为标准,就认定明清官话的基础音系也是北京音。显而易见,明清通行语"官话"原为明清"南京语"的专称,因此清末称北京语为"官话"须加"北京"限定。随着"北京官话"在使用中渐渐省略了限定词"北京"而摇身一变为"官话",专指南京语的"官话"则不得不加上"南京"以作区别,一些学者未谙其中关系而抬高北京语的历史地位,以致于中国近代官话史被扭曲。北京官话溯源亦为江淮官话在北方之后裔。朱棣先率淮兵扫北后又永乐北迁,大量江淮人迁至京津,致使京津成为江淮官话在北方的一个方言岛。更准确的界定是,现代北京话是以宋辽金元北迁中原正音为底层、以明代北迁江淮官话为主体,在清代大量满人换用汉语过程中阿尔泰化的一种方语。

明清官话以南京话为代表的江淮方言为基础,除了各种历史、经济、政治因素,还存在着浓郁的市民文学背景。京剧虽然在北京形成但并非土产,而是在徽、汉(湖北)两调合流的皮黄戏基础上发展起来的,并继承了昆曲中州韵的传统。近代百戏之祖昆曲源于南曲,而南曲韵宗《洪武》,即朱权为南戏所作的《琼

林雅韵》。《雅韵》其基础音系即明初南京话，与《洪武》大致相同[1]。京剧极力推崇的中州韵并非近代河南音，而是以江淮皖音和鄂音唱白，后北京演员以京音唱之，现代京剧音韵则成为一个混合音系。凡是京剧语音中与普通话声韵不同的上口字，就是仍保留在京剧唱念中的来自鄂、皖、苏的方音字[2]。

所谓"宋元以来，……有影响较大的小说，如《水浒传》、《西游记》、《儒林外史》、《红楼梦》等等。这些白话文学作品都是用北方方言写成的。"[3] 实际上，这些白话文学作品大多是用江淮官话写成，而江淮官话不是北方方言。施耐庵、吴承恩分别是江苏兴化和淮安人。吴敬梓是久寓南京的安徽全椒人。敬梓之子吴烺撰《五声反切正韵》，其基础音系即当时南京话[4]，又与江云樵合撰《金陵传声谱》。曹雪芹幼居南京，据研究者认为，《红楼梦》版本越早南京话成份越多，北京话成份增多盖传抄中改动所致。那宗训发现该书中曲韵阳声韵分押系统，合于南京、扬州一带方言，可以认为是曹氏母语的特色[5]。太田辰夫证明《红楼梦》语言以北京话为基础的例子"早起"，在江淮官话中为基本语词。至于太田辰夫论及《儿女英雄传》北京的土音土语所举例[6]，亦多见于江淮方言，如罕见字沍 u(焐)、硌 luo(摞)、礓礤 tɕiaŋ $ts^hA$(礓察)皆为江淮口语。今人从语言上考证《金瓶梅》，有山东方言说、徐州方言说、吴语方言说，皆各执一端。该书最

① 张竹梅《琼林雅韵研究》，宁夏人民出版社 1993 年。

② 杨振淇《京剧音韵知识》，中国戏剧出版社 1991 年。

③ 黄伯荣等《现代汉语》，甘肃人民出版社 1985 年。

④ 陈贵麟《杉亭集·五声反切正均音系探赜》，《语言研究》1994 年增刊。

⑤ 那宗训《从押韵看曹雪芹的语音》，台湾《大陆杂志》63:5，1981 年。

⑥ ［日］太田辰夫《红楼梦的语言》(1965)、《儿女英雄传的语言》(1974)，载《汉语史通考》

早刻于吴中（见沈德符《野获编》），与此书发现有关的沈德符（浙江嘉兴人）、徐文贞（江苏松江人）、刘承禧（湖北麻城人）、袁宏道弟兄（湖北公安人）、冯梦龙（江苏吴县人）多为江南及江淮官话区人。江淮地处豫鲁吴越之间，为南北语言交汇之地。宋元之际青徐流民多徙江淮，明初"洪武赶散"将苏州与杭嘉湖平原居民迁往江北，因此江淮方言中杂有北鲁南吴的俚语俗词。一些词语鲁、吴今已不用，《金瓶梅词典》[1] 中列入《难解词语待问篇》的"花靠"，至今江淮方言中仍用（含义"花样"）。因此《金瓶梅》虽然杂有原作者或修改者的方言俗语，但总体上是用明末官话写就，而不是一部北方方言小说。

总上所述，明清官话是宋元以降江南商业经济蓬勃发展、有着丰富的市民文化内涵，既运用于官场又为广大市民社会使用的，以保留在江淮方言中的中原正音为基础的通行语。换而言之，明清官话或明清两代汉民族共同语就是：以南京语音为标准音，以江淮话为基础方言，以通俗的明清白话小说为语法楷模的通行语。

## 五、七千年中国语历史演变的来龙去脉

在没有找到可靠的支点以前，远古语言的情况无法旁证推测。以中国境内三大考古文化系统的形成为座标，在这之前，东亚语言的状况假定为语网；在这之后，推定形成三大太古语区。公元前 3000 年左右，在三大太古语交汇之处，伴随着文明的接触、冲突和交融，原始华夏语混成发生。前 1300 年，以原始夷越语为主体的殷商华夏语具有文字系统。前 1000 年后，以原始氐羌语为主体的姜周语致使中原语言氐羌化，留下文献典籍。华

---

[1] 王利器主编《金瓶梅词典》，吉林文史出版社 1988 年。

夏语经秦汉统一演变为汉语。在游牧文明与农耕文明的持久冲突和交融之中，汉语主导语言的基础音系由西迁东又至南，在北留中原汉语的阿尔泰化与南迁中原汉语的夷越化同时，汉语在同化周边语言的过程中异化了自己。

在语言的历史长河中，中国语经历着生生不息的运动。华夏汉语的单词语音结构经历了双音→单音→双音的演变。最初的单词语音结构可能是双音节的，由于前一音节的元音弱化渐次脱落而变为复辅音音节结构，复辅音的单化使单词的语音结构单辅音音节化。单音节的辅音性韵尾又逐步失落，先丢失的是阴声韵塞尾，接着是入声的塞尾，最后是 -m 尾。对单词语音结构单音节化，复辅音单化及韵尾逐步失落所做的替偿，除了声调的形成、单元音的复化以外，还有合成构词法的广泛运用。合成词的产生表示原来的单词已弱化为语素，从而使汉语的非根词性单词以新的方式重返双音节结构。华夏汉语的声调并非与生俱有。当单词语音结构是双音节时，华夏汉语是非声调语言；当单词语音结构是复辅音音节并带辅音尾时，出现伴随性声调；在单词语音结构单辅音化和辅音尾失落过程中，汉语形成音位性声调。随着双音节合成词的大量产生及阿尔泰语的影响，声调由多渐少。北方汉语中的一些字调弱化为轻声，显露出声调逐步消失的端倪。形成了华夏汉语：无声调→伴随性声调→音位性声调→声调渐少→弱化调的演变轨迹。

华夏汉语的句法结构也经历着巨大嬗变。粘着型或不完全屈折型中的语序并不是刚性手段。随着形态失落向孤立型方向发展，一些附加成分或实词虚化为虚字，语序和虚字上升为主要手段。东亚北方语的是 SOV 及 AN 逆行结构，东亚南方语的是 SVO 及 NA 顺行结构。周秦时期南方型与北方型共存，两汉以后逐步稳定为 SVO(南方型)与 AN(北方型)，这种特殊的句法

结构也表明了汉语的混成性质。在 NA 转变为 AN 的过程中，产生了类别词[1]。唐宋以来，北方汉语受阿尔泰语的影响出现了 S(把)OV，使近现代汉语成为一种以 SVO 型为主兼有 SOV 型特点的语言。语言的多向性、层叠性接触、换用和混成，促成华夏汉语形态的失落而蜕变为孤立结构。与西方进化论者所持有的孤立型→粘着型→屈折型演变相反，东亚大陆语言的发展史实表明，孤立型才是形态语言演变的结果。

南亚—南岛语源于原始南方语，南岛居民是公元前 4000 年从大陆南部开始南迁入海的夷越人。现在的阿尔泰诸语言是原始胡狄语的后裔。在游牧和农耕文明的冲突交融中，进入中原的阿尔泰人主动汉化并换用了汉语。现在的藏缅诸语言是原始氐羌语的后代。汉藏词汇同源是因为姜周使华夏语氐羌化，而藏语句法更像阿尔泰语，是因为保留着太古氐羌语的粘着结构。朝鲜半岛—辽东半岛—山东半岛—江淮平原—华南地区，在远古同属水耕文化，但是朝鲜半岛的北面是阿尔泰世界，由此形成了底层南岛语而表层阿尔泰语的、又向汉语大量借词的朝鲜语。海岛日本的最早居民或是来自大陆的夷越人，或是从南渡海而来的马来人，日本语当属太古夷越语的一支后裔。处于汉语和南岛—南亚语之间的侗台—苗瑶语，一些语言中还保留着粘附形式，但绝大多数语言是单音节化、有声调、以语序和虚词为主要手段的孤立结构，或许这些语言从原始华夏语中分化出来时已处于孤立化过程中，或许汉语通过文化辐射或借词影响致使这些语言趋于孤立化。

一些学者将见于汉语书面语稍晚而存在于南亚语中的词，定为上古汉语中早已完全"归化"的外来语成分，并由此推定早

---

[1] 洪波《汉语类别词起源初探》，《现代语言学》，延边大学出版社 1990 年。

期台语对汉语产生过巨大影响[1]。然而，无论是南岛—南亚语、侗台—苗瑶语、还是华夏汉语，其动脉中都流动着太古夷越语的血液，汉语书面语稍晚出的词从“先祖”那里继承下来的固有词，而并非从“弟兄”那里借来的词。之所以确定见于南方语中的词是汉语中的外来成分，就是因为首先预设了“汉语仅与藏缅语同源”。因此，在“华夏汉语混成发生论”的基础上，我进而提出“东亚南洋语言文化圈假说”，力图为未来的汉语语言史、中国语言史和东亚南洋诸语言的研究提供一个新的理论模型。

## 第三节　东亚南洋语言历史文化圈假说

尽管从东西方古老的传统经济方式的差别，就可以窥见东西方语言关系类型的迥然不同，但是由于历史比较语言学发轫于19世纪的西方，并且依据世界语言中少见的、由史前游牧文明所决定的分化轨迹的印欧语事实建立了谱系树模式，因此当西方学者和那些受西方语系说熏陶的东方学者以“印欧语眼光”审视远东语言的时候，谱系树模式与远东语事实的龃龉必然导致东亚南洋这一广袤区域的语言系属关系的众说纷纭[2]。如果把东亚南洋语言文化圈的语言关系的早期类型确定为非语系的语网和语区，印欧谱系树眼光和东亚南洋语言事实龃龉所导致的语言系属关系的歧异，也许有希望迎刃而解。

---

① P.K. Benedict: *Early Chinese Borrowings, Genetic Relationship, Diffusion, and Typological Similarities of East and Southeast Asian Languages*, 1976.

② 游汝杰《中国语言系属研究述评》，《云梦学刊》1996年第3期。

## 一、东亚南洋语言关系假说的歧异

(一)汉藏—印欧语同源说

1871和1872年,远东语言学研究的先驱艾约瑟(J. Hdkins)在《中国在语言学中的位置:揭开欧洲和东亚语言具有共同来源的一个尝试》中和史莱格尔(G. Schlegel)在《汉语—雅利安语》中相继提出这一欧亚语言同源假说。1963和1965年,美国谢飞(R. Shafer)的《欧亚语》和《欧亚语言学的超级语系》中发扬了这一同源说。1995年,加拿大蒲立本《汉语的史后和史前的联系》中,从语音结构和同源词两方面试图论证汉藏语和印欧语的发生学关系。他所构拟的上古汉语的两个元音的系统,与原始印欧语元音系统存在对应关系。在辅音方面也存在一些对应规律,如上古汉语 *ŋ- 对应原始印欧语的 *g-,并考证出24个同源词。蒲立本并且认为汉藏语和印欧语的分化在6000年以前,即在利用历史比较法构拟出来的原始印欧语之前。德国学者张聪东在《古汉语复辅音在印欧语词汇的印证》(1991)中,从谐声字和印欧语词根音义相通的"同源系列"入手,试图揭示汉语和印欧语的相似性。如果这些同源系列对应可靠,那么东亚语和印欧语之间的古老的史前同源关系,其背后隐含着意义更为深远的人类语言起源单源论的证据。

(二)汉语—藏缅语—苗瑶语—侗台语同源说

关于汉藏语言的历史比较研究,最早的论著是莱顿(B. J. Leyden)在1808年发表的《论印支诸民族的语言和文学》。1881年和1887年,格鲁布(W. Grube)和拉古勃里(T. de Lacouperie)先后在《汉语的历史地位》和《汉文创制前的汉语》中讨论汉藏语言的系属问题。1911年,科诺(S. Kenow)在《印度语

言调查》中最早提出侗台语属于汉藏语系，同时认为藏缅语的词序为 SOV，汉语和台语的词序为 SVO。在对藏缅语所知甚少的情况下，法国马伯乐(H. Maspero)提出了汉藏语系分为东支汉—台语族和西支藏—缅语族的设想。这以后，施米特(W. Schmidt，1914)、特罗姆贝提(A. Trombetti，1923)和普鲁伊路斯基(J. Przyluski，1924)都提到汉藏语或印支语。

1934 年，赵元任《中国分省新图》中在阿尔泰语类之外，分出包括汉语系、太语系、苗瑶语系和藏缅语系的汉藏语类。1937 年，李方桂以当时的语言调查和理论为基础，在《中国年鉴·中国的语言和方言》中确定了包括汉语语族、侗台语族、藏缅语族和苗瑶语族的汉藏语系，认定汉语和侗台语同源的主要依据是单音节和声调。张琨(1947)又证明了苗瑶语声调和汉语的一致性(1970)，但后来马提索夫(J. Matisoff)等提出"声调发生"的概念，认为声调的发生和语言发生学无关。邢公畹(1948)和俞敏(1949)从语法角度论证汉藏同源，但 60 和 70 年代，河野和桥本万太郎认为汉语和藏语的语法系统差别太大，可以说是两种不同的语言结构类型。尽管如此，李氏四分系统一直是中国语言学界长期认同的传统分类法。虽然《中国语言学报》(1972)在重新刊登李氏旧文时并加上了"李氏的论述在本质上仍然是正确的和有用途的"编者按，但是李方桂在《台语》(1974)中却承认台语和汉藏语系的关系从未明确建立。陈保亚认为，与美国学者的同源标准不同，中国学者的"汉藏语系说"研究方法采取的是同构标准[2]。然而这仅仅是表象，在一些中国学者的潜意识中是"传统疆域标准"——即宁可把境外语言纳入汉藏语系，而绝

[2] 陈保亚《百年来汉藏语系谱系研究的理论进展》，《语言学论丛》第 21 辑，商务印书馆 1998 年

不可以将境内貌似汉藏语的语言认定为非汉藏语系,除非这种语言和汉藏语的差距实在太明显。

(三)台语—加岱语—印尼语联盟说

与中国学者的"疆域文化决定论"不同,美国学者采取的是"语言文化决定论"。从1935年,美国的一些语言学家在人类学家克罗伯(A.L. Kroeber)的主持下,开始执行汉藏语言比较研究计划。1942年,白保罗的《台语—加岱语—印尼语:东南亚的一个新联盟》,提出台语(相当于侗台语族中的壮傣语支)、加岱语(包括黎语、拉嘉语、仡佬语等)与印尼语才具有发生学上的关系,而不是与汉藏语,台语和加岱语只不过在汉语影响下发生了很大变化。

白保罗认为:1.虽然加岱语[1]是单音节、孤立型,和台语一样有完整的声调系统,但声调和词音单音节结构不能作为语言亲属分类的依据,而台语和汉语的对应词是借词。2.加岱语和台语、印尼语一样,是NA、VO的顺行结构。藏缅语的原始形式是逆行结构OV,而卡伦语和汉语的VO顺行结构是受澳泰语、苗瑶语、孟高棉语影响的结果。3.台—加岱—印尼构成一个单一的语言连续体。虽然加岱语和台语的单音节与印尼语的多音节不同,但这是加岱语和台语从汉语中借来声调且音节脱落的结果。尽管加岱语更接近台语而不是印尼语,然而加岱语是这三位一体的过渡成员。4.台语—加岱语—印尼语的祖先是原始澳泰人,也许历史上的澳泰人文化技术比北方汉人发达。台语词汇之所以与印尼语相差甚远,而与汉语词汇近似,是因为上古汉语向澳泰语借词后又贷词给澳泰语的结果。5.加岱人是中国中南部的土著居民。仡佬人的老家在贵州,由于苗族的进入

[1] "加"Ka是拉嘉语、仡佬语"人"的前缀,"岱"Dai是黎族自称。

而迫使他们南迁。白保罗进一步推断，长江流域以南是玻里尼西亚人或原始澳泰人的居住地，他们从中国南部海岸经过海南岛迁徙，北支迁移到台湾，东支迁移到菲律宾，南支迁移到越南、婆罗洲、爪哇、苏门答腊和马来半岛。印尼语曾分布在南大陆的语言底层，可以从加岱语中看出来。6.在孟高棉语为主体的地域以南的马来语今天异乎寻常的位置，只能解释为它是从印尼诸岛由海路迁徙而来的。北面的台语—加岱语、东面的占语和南面的马来语，在孟高棉语的周边呈现出弧形分布状态。而苗瑶语可能是原始南方语的一个分支。

台语—加岱语—印尼语的原型，就是以学识渊博而著称的奥地利神父 W·施米特《世界的语系和语言圈》所划出的原始南方超级语系(Austric Superstock)的北支，与南支孟高棉语的分裂一定发生在华南—印度支那地区，随后这两支分别在南北定居下来。德国史图博(H. Stubl)在《海南岛的黎族：为华南民族学研究而作》(1937)中指出，印尼人和密克罗尼西亚人许多相同之点。黎族原有的印尼文化特征早已为大陆上汉化的加岱族所忘却，台—加岱—印尼共同的人种图象早已磨损得无从修复。

与白保罗共同完成《汉藏语言学》手稿的谢飞，早在 1938 年就通过基本词汇比较，发现台语和汉藏语之间缺少语音和语义上的紧密关系，认为台语不属于汉藏语系。虽然在《汉藏语系语言的分类》(1955)中仍然将台语归入汉藏语系，但认为汉语与藏语的关系比与台语的关系更接近。即使台语与汉藏语有关系，那种亲属关系也很遥远。同时，谢飞的汉藏语系说却把黎语和苗瑶语排除在外。奥德里古尔在《台语的语音和词汇》(1948)中认为，台语和汉语接近的常用词是数词、军事术语和工艺术语等一些容易借入的文化词汇，与之相反，人体部位名称和农业词语都和相应的汉语很少具有共同之处。通过基本词语比较，奥德

里古尔也得出了马来一玻里尼西亚诸语言发源于亚洲的结论。伊万诺娃(E.B.Иванова)在《论台语及其在发生学分类法中的地位》(1970)中认为:台语民族的祖先属于包括现代印尼和孟高棉诸民族祖先在内的同一原始人种集团。在起源上大概与汉藏诸民族没有关系,但是从太古时代就开始接触了,并且这些接触在台语族的语言文化中留下了若干痕迹[1]。赞同这一假说的还有易嘉乐(S. Egerod)、福雷斯特(R. Forrest)、包拟古(B. Nicholas)和马提索夫等。

与白保罗的台语—加岱语—印尼语联盟说得到国际语言学家的普遍赞同相反,绝大多数中国语言学家还难以接受侗台语和苗瑶语从汉藏语系中分离出去的结论,依然固守传统四分法。尽管如此,还是有一些中国语言学家依据语言事实,从"疆域文化决定论"中突围出来。闻宥在《台语和汉语》(1957)中通过台语和汉语基本词汇的比较,认为二者没有同源关系。罗美珍在《试论台语的系属问题》(1983)中认为,台语是在汉藏母语基础上发展而来的一种语言。古代台语民族即百越族,其语言原是多音节型。《世本·居篇》:"吴蛮夷言多发声,数语共成一言",扬雄《方言》所载的吴越词语是双音节,如怜职(爱)、煦暇(热)。作为百越的一支,南迁以后的印尼人保留了古老的粘着型和多音节。留居大陆的台语先民则放弃了百越语或原始马来语,换用了原始汉藏语。原始汉藏语分化以后,独立发展成一支语言的台语残留着原始马来语的底层。倪大白在《中国的壮侗语与南岛语》(1988)和《南岛语与百越诸语的关系》(1994)中,把侗台语和印尼语、阿眉斯语、布嫩语、回辉语、菲律宾语和马来语进行比较,得出侗台语和南岛诸语言的基本词汇具有一致性的结论,并

[1] [苏俄]伊万诺娃《论台语及其在发生学分类法中的地位》。

且参照考古学和人文史资料，证实侗台语和南岛诸语言同出一源，但侗台语的语言类型已经转变，侗台语的系属应称之为"马来—支那语"。这些研究表明从基本词汇角度看侗台语和印尼语有亲属关系，从结构类型角度看侗台语和印尼语有亲属关系。根据混合语理论，侗台语可能是一种历史混合语（从公元3世纪开始），侗台语的"基础来源语"是南岛语，"混入来源语"是汉藏语。

雅洪托夫（С. Е. Яхонтов）在《语言年代学和汉藏语系诸语言》（1964）中认为，单一的加岱语支是不存在的，加岱诸语言互相接近的程度和台语支诸语言一样，而不是更接近些。无论是台语支语言，还是加岱语支语言，只能在斯瓦迪斯词表中找到10～15个与现代印尼语一致的词语，这表明台语—加岱语与印尼语之间的亲属关系比台语和加岱语之间，甚至比汉语和藏缅语族之间的关系都要远得多。雅洪托夫的语言年代学研究结果和白保罗的分类法一样，都认为台语和苗瑶语不属于汉藏语系。语言年代学方法以语言分化模式为前提，排除了诸语言之间的接触或混合，只是封闭性语言匀时发展的一种比较方式。因此必须折算或扣除接触或混合所引起的加剧的词衰减变量，以弥补均时语言年代学方法之严重不足。如果考虑到汉语对台语的巨大影响，就不能够仅仅依据台语与印尼语的一致性词语的10～15个来计算分化年代，而应当进行年代矫正。

（四）澳泰语系说

在《台语—加岱语—印尼语：东南亚的一个新联盟》中，白保罗以"新联盟"为基础在《澳泰：语言和文化》（1975）中进一步提出澳泰语系说。1943年写成的《汉藏语言概论》，由马提索夫找

出来详加注释,直至 1972 年得以出版[①],公布了白保罗将侗台语和苗瑶语划出去的汉藏语系分类法。

<table>
<tr><th colspan="6">原始南方语(Proto Austric)</th><th colspan="3">原始汉藏语</th></tr>
<tr><td colspan="4">北支澳泰语系</td><td colspan="2">南支孟高棉语系</td><td rowspan="2">汉语</td><td rowspan="2">藏缅语</td><td rowspan="2">卡伦语</td></tr>
<tr><td>台语</td><td>加岱</td><td>印尼</td><td>苗瑶</td><td>孟高棉</td><td>越南</td></tr>
</table>

雅洪托夫在《语言年代学和汉藏语系诸语言》(1964)中用语言年代学方法,提出与白保罗相似的澳斯特利(Austric)语系说。

<table>
<tr><th colspan="6">原始澳斯特利语</th><th colspan="7">原始汉藏语</th></tr>
<tr><td rowspan="2">印尼</td><td colspan="3">侗台</td><td colspan="2">孟高棉—苗瑶</td><td rowspan="2">汉语</td><td colspan="5">藏缅</td><td rowspan="2">喜马拉雅</td></tr>
<tr><td>黎语</td><td>侗水</td><td>台语</td><td>孟高棉</td><td>苗瑶</td><td>藏语</td><td>景颇</td><td>缅语</td><td>彝语</td><td>卡伦</td></tr>
</table>

雅洪托夫认为:统计证明,在苗瑶语中可以找到很多与孟高棉语共同的词,而与侗台语共同的词却很少。苗瑶语也属于南方语系,而且在这一语系中比其他所有语言更接近于孟高棉语族,这可能是因为苗瑶语脱离孟高棉语族的时间,比南方语系分化为三个语族的时间要晚。

(五)汉语—南岛语同源论

1706 年荷兰 H·莱兰特发现了南岛诸语言的亲属关系,1836 年洪堡特用马来—波利尼西亚语这个术语来称说这一语

① 白保罗《汉藏语言概论》(乐赛月、罗美珍译),中国社会科学院民族研究所语言室印行

系的语言。"南岛语系"这一术语是 W·施米特在 19 世纪末提出的,它包括印度尼西亚、美拉尼西亚和玻里尼西亚诸语族的语言。关于南岛语和汉藏语的关系,20 世纪初就引起语言学家的注意。1916 年和 1922 年,孔好古发表了《澳斯特利语言和印度支那语言间值得注意的关系》和《澳斯特利语—印度支那语新的对应》。W·施米特用"共同底层"建立起包括这些语言的印支(汉藏语)—南方(南岛—南亚)超级语系。1942 年,吴克德发表《论马来—玻里尼西亚语言和印度支那语言的关系》。通过泰语和南岛语、藏缅语和南岛语的词汇对比表,证明在澳斯特利语(南亚语和南岛语)和汉藏语(汉泰语和藏缅语)之间具有发生学关系。值得注意的是,与藏缅语相比,汉泰语和南岛语的关系更为密切。

1990 年,法国沙加尔论证了汉语南岛语同源说。依据白保罗的台语单音节是音节脱落结果的观点,认为汉语的单音节词和南岛语的多音节词的末一音节对应。汉语的声调和南岛语的辅音韵尾对应:原始南岛语的 * -s 韵尾和上古汉语去声 * -agh、* -idh、* -adh 对应;原始南岛语的 * -p、* -t、* -k 的和上古汉语入声 * -p、* -t、* -k 对应。汉语辅音聚合关系和南岛语的对应:上古汉语的不送气清塞音、送气清塞音、浊塞音的对立,相当于原始南岛语的清塞音、鼻冠清塞音、浊塞音的对立。汉语介音与南岛语中缀的相近:上古汉语的介音 * -j-和 * -r-是具有构词作用的中缀,和原始南岛语中缀 * -in-、* -ar-相近。已发现的汉语和南岛语关系词有 222 个,其中 17 个见于斯瓦迪斯词表。

谢飞(1955)曾提出:汉藏语和 Athapqskan 语配比的词根,与汉藏语和南亚语配比的词根完全不同,这是否表明汉藏语系是否实际上是两种更原始的语系:太平洋语系和南亚语系的某种混合物呢? 伊万诺娃(1970)认为把汉藏语系、孟高棉语系和

马来—玻里尼西亚语系联合成一个太平洋语系的理论,现在既没有可能创立但也不必拒绝。在汉语南岛语同源论提出以后,邢公畹(1991)和郑张尚芳随即表示支持,提出了包括汉语、藏缅语、侗台语、苗瑶语、南岛语和南亚语的华澳语系说。

(六)汉语—藏缅语—高加索语同源说

汉语—藏缅语—北高加索语同源说,也称之为汉—蒙语系说。1853年,霍奇森(B. H. Hodgson)在《论汉藏语和蒙古语的亲属关系》中,对汉藏语和高加索语亲属关系的论证做过尝试。1939年,波达(K. Bouda)的《苏美尔语与巴斯克语、北高加索语以及藏语的关系》一文继续探讨了这一问题,因此西方的一些学者认为存在一个"汉—蒙语系"。1995年,俄国学者斯大罗丝汀(S. Starostin)在《上古汉语词表:一次历史的透视》中通过35个基本词汇的比较,得出汉语和藏缅语24个同源、和北高加索语13个同源、和南岛语4个同源的结果,推断汉语和藏缅语关系最密切,其次是与北高加索—叶尼塞语。斯大罗丝汀提出南岛语和台语的关系相当于汉藏语和北高加索语的关系,因此汉语南岛语同源说必须放弃。

(七)苗瑶语的系属歧异

苗瑶语系属大致有三种假说。1.孟高棉语族说。英国戴维斯(H. R. Dawies)在《云南:印度和扬子江的链环》(1909)中提出此说。他还认为孟高棉语(不包括蒙达语的南亚语)、掸语、汉语和藏缅语有共同的起源。2.澳泰语系说。白保罗(1942)认为苗瑶语属于澳泰语系,但有所存疑。与之相似的是雅洪托夫(1964)认为苗瑶语与孟高棉语族关系密切,属于澳斯特利语系。3.汉藏语系说。王辅世在《苗瑶语的系属问题初探》(1986)中,批评了戴维斯和白保罗的缺陷,从同源词和类型学的相似性以确定苗瑶语属于汉藏语系,但对之属于什么语族或是否独立为

语族却未明确。

（八）白语的系属歧异

白语的系属大致有五种假说。1.孟高棉语族说。戴维斯在《云南:印度和扬子江的链环》中通过对一百个词的分析，得出汉语源42个、藏缅语源33个、孟高棉语源23个、掸语源2个。依据现代白语地区没有孟高棉语，推定这23个必定是白语的原词。2.混合语说。李澡在《大理县志稿·方言》(1916)中认为白语是固有蛮音从宋元以后与彝语混合，又受到明初吴越一带迁来的汉族居民语言影响而形成的混合语。罗常培的《语言学在云南》也主张白语为"夷汉混合语"说。徐嘉瑞《民家语新诂》中认为白语是汉语和白蛮语的混合，当归入汉藏语系而别列语族。徐将"民家"考释为"名家"，其实"民"是汉藏人的自称。3.藏缅语族说。闻宥《民家语中同义词之研究》(1940)批评了戴维斯的白语一孟高棉语族说，认为白语中最古的成分是藏缅语，至于孟高棉语的成分简直没有。李方桂早年(1937)曾经认为白语属于彝语支，但晚年(1980)认为白语受到汉语的强烈影响，系属值得怀疑。周耀文的《略论白语的系属问题》(1978)认为，白语是属于藏缅语族的一个单一语支。赵衍荪在《白语的系属问题》(1982)中同意白语属于彝语支。4.上古汉语方言说。三十年代，白族学者赵式铭父《白文考》认为，白语中保留大量周秦汉语"雅言"词汇，属于在汉代以前就形成的汉语方言，后来受到与之杂处的其他民族语言的影响成为白语。张海秋认为白族的"言语声音，实能保存三代以来之古语古音"。5.汉白语族说。白保罗(1942)的汉藏语分类把白语归入汉语族。郑张尚芳在《汉语和亲属语同源词根词及附缀成分比较上的择对问题》(1995)中认为白语属于汉白语族，在《白语是汉白语族的一支独立语言》(1999)中论证了"汉白语族"假说。吴安其在《藏缅语的分类和

白语的归属》(2000)中认为，白语在语音、语法和核心词的构成上与彝缅语比较接近，词汇上与汉语的相近是借用造成的，从语音和语序的历史看，白语应是彝缅语的一支。实行古氐羌人习俗火葬的白族这一藏缅支系到达滇西北可能是战国至汉，其习俗不同于分布在洱海较早从四川进入云南的古藏缅"昆明人"和滇池一带的"滇人"。既然白族远祖是古氐羌人的一支，那么与汉语之间的亲缘关系就是存在的。问题是要使"汉白语族"成立，就必须证明白语是什么时期从华夏语中分化出来的。

（九）阿尔泰假说和阿尔泰语系

阿尔泰假说问题，可以说就是阿尔泰语系假说问题。迄今为止，围绕阿尔泰假说仍然有许多争议，其焦点就在于是否承认阿尔泰语系的存在，或者说阿尔泰诸语之间是亲属还是借贷关系。第一个注意到突厥、蒙古、通古斯语言之间存在相似性的是瑞典冯·斯特来林贝尔格(Von Stralingberg)，在18世纪上半叶提出分为六个语群的"塔塔尔诸语言"假说。一个世纪以后，拉斯克(R. Rask)重新命名为斯泰基诸语言，并将其中的语言做了调整。由于阿尔泰诸语言是粘着结构，因而这一特征一度成为判断是否与之相关的语言的主要原则。马克斯·缪勒(F. M. Müller)将暹罗语、藏语、拉德维达语、马来语也扩大进来，称之为"都兰诸语言"。M·A·卡斯特林(M. A. Kastling)认为语言同源的本质是词素的一致性，只有乌戈尔—芬兰语、萨莫耶德语、突厥语、蒙古语、通古斯—满语才构成相关语群，现通常分为乌拉尔语群和阿尔泰语群。兰铁司(G. J. Ramstedt)在《阿尔泰语言学导论》中，确立了原始朝鲜语、原始突厥语、原始蒙古语、原始通古斯—满语的阿尔泰语系四种方言说，并建立了许多语音对应关系。

不接受阿尔泰假说的学者主要有两类：一类学者并不否认

阿尔泰假说，但认为证据尚不充分；一类学者认为阿尔泰假说不能成立，把阿尔泰语中的这些共同成分解释为接触的结果。虽然阿尔泰假说的理论尚不完善，但迄今为止，反对者并没有提出使语言学家能够接受的、足以否定这一假说的充分证据。之所以有学者不接受阿尔泰假说，一方面，是因为谱系树模式或“语系”说的禁锢。以严格的印欧语系眼光来看世界语言，符合“语系”的除了印欧语系，可能就是闪含语系。狭隘的“语系”说禁锢了欧洲学者的眼界，认识不到——不同的生态环境和经济方式决定了语言的发展方式。在早期语网的基础上逐步分化，又随着游牧经济相互借贷词语的阿尔泰诸语言，确实不可能也无须具有印欧语系那么严格的词语对应规则和明显的分化层次。以游牧经济为基础的语言，结构类型的相似性应是衡量这类语言的亲缘关系的标准之一。依据元音和谐规则、粘着结构类型和部分基本词语对应，可以认为阿尔泰诸语言是以原始语网和相互借贷词语为基础的亲属语言。至于阿尔泰诸语言的内部关系，比如原始突厥语和原始蒙古语的关系、突厥人何时迁至西域、突厥人和西域土著的关系等，还值得深入探讨。一方面，西域土著是古雅利安人种，当地曾经存在过印欧语系的几种语言或方言；另一方面，与蒙古人的草原和丛林文化有所不同，突厥人是西域地区的“绿洲文化”。

（十）东亚南洋语言文化圈假说

1990年，李葆嘉提出“华夏汉语混成发生论”。稍后，在《中国语的历史和历史的中国语》(1996)中进一步提出“东亚南洋语言文化圈假说”，这一语言文化圈包括澳台语(日本语、朝鲜语、南岛语、侗台语)、南亚语(南亚语、越南语、苗瑶语)、汉语(南方汉语、北方汉语)、藏缅语(彝语、藏语、缅语、景颇语、羌语)和阿尔泰语(通古斯语、蒙古语、突厥语)。又在《语言关系类型学：对

谱系树模式的超越》(2000)中提出涵盖谱系树模式的世界语言关系类型的四分法。

## 二、东亚南洋语言文化圈的史前渊源

根据近40年来的考古学成就,中国境内的旧石器遗存和新石器文化的中心并非仅在黄河流域,反之,多种考古文化的冲突交汇才形成了后来居上的中原文化。虽然考古学界最近宣布,1997年在河北南庄头出土的陶片、石器、角器以及谷物这些一万年前的新石器文化遗存,东亚制陶史的开端提前了3000年,但仍然没有证据说明这就是东亚新石器文化的原始母体,同期的考古文化就有江南的溧水神仙洞遗存。汉藏南岛综合体或单一的新石器文化母体可能并不存在,东亚南洋新石器文化母体更有可能存在长江流域和黄河流域两个源头。在讨论了若干预设的基础上,才有可能探索东亚语言的史前渊源,以建构东亚南洋语言文化圈假说。

### (一)东亚南洋语言文化圈假说的若干预设

#### 1.夏娃理论和蒙古人种的形成

依据生物分子学的研究和考古文化的印证,现代人或新人起源于距今15~12万年之际的东非大草原。欧洲科学家依据线粒体DNA特征研究的分组结果,推测现代欧洲人的99%是由大约在4.5万年之际先后来到欧洲的七位女性祖先繁衍而来,并提出其"白化"在距今3万年之际。由此可以旁推,蒙古利亚人种也应形成于距今3万年前,东亚大陆的新人应是在这之前从中东迁来。

#### 2.采集狩猎经济型的连锁语网

托尔斯托夫曾经提出"原始语言连锁扩散性假说"根据澳大利亚土著语之间很难找出明确界限等事实,可以推定新旧石器

时代之交为"原始语网"。由于新石器革命进程的不平衡和生态环境所导致的自然经济方式的不同，致使世界语言逐步形成游牧型语系、农耕型语区、采集狩猎语网和渔猎型语珠等四种关系类型。在东亚大陆南部、东南亚以及南岛这一广袤区域，直至公元以后，一些地区还处于新石器时代早期，甚至停留在旧石器，与印欧语的语系式分布状态不同，语网关系成为这一区域的特征。欧洲学者并非没有觉察到这一状况，本维尼斯特认为：在远东"存在着从藏语到缅甸语，再到怒江语群诸语言，其后到孟高棉语，最后到大洋州诸语言的逐步过渡。这些语群中的每一语群都不可能界定得清晰，每一语群都有其特点。其中一些特点使它与前一语群衔接，而另一些特点则使它与后一个语群衔接……植物学家很熟悉这种'连锁血缘'，也许只有这种类型的分类法才是我们现在构拟的这一范围内的一些大语群的惟一合适的分类方法。"[①] 面对这一区域的复杂性，新石器时代早期的原始南方语仍然是原始语网。

3. 语言亲缘关系的历史性质

正如安德罗诺夫所说："从语言本质的自然主义理解，留下的所谓一个语系内部诸语言的亲属关系不可能永久性的和最初性的，而是带有历史性的——它是逐渐形成的，同样也能逐渐衰弱以致消失。"[②] 语言亲属关系的历史性观点揭示了语言关系的动态性，从而指明了语言亲属关系的认定只能以某一时期内的存在为依据。从理论上说，可以把这一上限确定为新石器时代的开端但实难统一，因为原始群落进入新石器时代的先后不同，借以推测早期语言同源或接触的新石器文化遗存的分布并不平衡。依据东亚大陆的考古文化系统，一般只能追溯到距今7000

①② 转引自伊万诺娃《论台语及其在发生学分类法中的地位》。

年左右。

马伯乐认为:"古代汉族文化与泰族、彝族、苗族文化之间有那么多的共同之点和显著的相似现象,以致令人不由自主地产生了所有这些民族在史前时期就存在共同文化的想法。有过这样的文化共同体是确定无疑的,它可以用汉、藏、彝、暹罗等语言中的大量亲缘关系词得到证实,比方说在所有这些语言中都有共同的数词。虽然不能根据这些相似形式就遽然断言,这些语言一定是亲属语言,但是这些词语可以表明:这些民族的语言和文化在史前时期就已经彼此发生了深刻的影响。也许它们不存在语言的亲属关系,可是明显地具有文化的亲属关系。"[①] 既然这些民族的语言文化在史前时期就彼此发生了深刻影响,因此如果限定在某一个阶段,它们之间就具有某种程度的亲缘关系。在语言亲缘关系的长河中,存在可追溯的"同源亲缘",也存在本非同源的"接触亲缘";反之,存在本非亲缘的"真性异源",也难免亲缘泯灭的"假性异源"。

**4.语言亲缘标准的综合性质**

对于东亚地区语言结构的相似性,一些欧洲学者强调不是起源的共同性,而是接触渗透的趋同性。奥德里古尔(1961)认为:"在结构中的相似性足以论证语言在发生学上的亲属关系时,认为有可能把东南亚语言地图的复杂性简化为二分法,使汉藏诸语言与南亚诸语言相对立……吴克德曾经试图使这些语言的关系互相靠拢。现在既然强调语言邻接的意义及其最终的表现——双重语言制,就不那么倾向于否定结构借用的可能性了,用二分法或单一性简化语言地图是不能令人信服的……按照这

① 转引自伊万诺娃《论台语及其在发生学分类法中的地位》。

一假设[1]，除明显的现象以外，东南亚各种不同的语言区域在发生学上彼此间没有联系。"[2] 此外，张琨（1969）、马提索夫（1973）和白保罗（1976）等，也同样认为类型学标准，特别是声调不足以证明同源关系。

也许，因为汉语的声调也可能是后起的，因此声调和词音单音节性确实不能作为同源的证据。尽管受汉语的影响，侗台语和苗瑶语等语言的结构类型发生了转换，但是仍然不能以语言接触中的"结构借用可能性"否认发生学上的"语言结构相似性"。如果史前语言的亲缘关系和结构类型不具有一致性，那么就无法讨论语言的同源。这一亲缘关系和结构类型的一致性，至今在许多关系语言之间还可以看到。所谓语言的同源，包括根词同源和结构同源两个方面。依据本维尼斯特"在史前时期就已经按照类型相似组成语言集团的话，那么从历史的观点来看，我们觉得类型的相似性似乎是发生学上的亲属关系的一个证据"[3]的观点，那么与他另提出的"语言结构类似不能成为语言同源证据"不免互相矛盾，也就是说后者不具备历史语言学理论的普遍意义。

无论像中国学者为维护"汉藏语系四分说"所坚持的语言结构的稳定性（同构标准），还是像外国学者坚持把基本词汇的对应作为惟一标准（同源标准），其实都是无的放矢。在语言接触过程中，语言的任何要素从来不存在绝对的稳定性，问题在于接触的方式和深度。虽然结构类型的相似性可能出于长期影响，但基本词汇也并非不可能大量借用。依据语言的无界接触观，语言中的任何成分都可以渗透和借贷，不但结构类型和音系格

① 根据上下文是指"区域接触性借用假说"。

②③ 转引自伊万诺娃《论台语及其在发生学分类法中的地位》。

局可以借贷,就是核心词汇也可以借贷。西方历史比较语言学所标榜的"语音对应规则",除了可以作为语言同源的证据,也可能是语言趋同的结果。语言接触中最活跃的要素词语,其语音结构特征会对接触语言的个别音位乃至音系格局发生影响,其语义结构方式会对接触语言的语法范畴乃至结构类型发生影响。随着语言接触的深入,语言的任何要素都可能发生变化即丧失其稳定性。因此语言亲缘关系的认定标准,应当是音系格局、结构类型、基本词汇以及文化特征、历史追溯的综合标准。

**5. 同源词和借源词的历时相对性**

词语的相似性有同源和借用两种可能。所谓"同源",应当包括新石器时期以来的语系分化和旧石器时代的语网延展两种情况下的同源。虽然连锁型语网关系是东亚南洋某些语言区域至今仍然可见的基本特征,但是不可否认新石器时代强势文明的形成促使某些地区的连锁型语网向同化型语区方向的演变趋势。随着强势文明的形成发展,其语言文化因子必然会向周边扩散,从而促使接触语言或少或多的同化。由于东亚地区语网关系和文化扩散的交叉性,致使同源词和借贷词的区别显得非常复杂,不得不以"关系词"来指代"语词的相似性"。但是前代的借用可能成为后代的同源,二者只是不同历史层次上的关系词,因此相互之间没有绝对的界限。由于年代久远的早期借用和可能同源之间的界限趋于泯灭。当借贷到达一定程度,双方也就成为亲属语言。最全面的借用也就是语言的换用。较长历史时期的持续借用和较短历史时期的迅速换用,只是过程长短不同而实质无别。因此无论是借用和换用,还是渗透和混合,都没有截然的界限。

**6. 语言混合是深刻接触的自然结果**

以往的中国语言学界,谈到语言融合也只是所谓同化,或斯

大林主张的胜利者的“吃掉”，而根本不考虑通过混合会产生一种新的语言。在人类历史上，语言的混合是若干语言长期接触的必然结果。语言接触必然导致语言要素的渗透以致混合，而对于混合过程中的语言来说根本不存在什么最稳定的要素。在渗透较弱和层次清楚的两种语言接触的情况下，可以根据其主流特征确定其归属关系；在混合深刻和层次复杂的多种语言接触的情况下，惟一明智的选择就是确定为混合语，揭示其“基础源语”和“混入源语”。对于这种混合语，到底是新语种的产生还是原语言的演变，这不是一个语言学问题而是一个民族学或文化学的问题。语种的质的规定性，绝不等于语言的差异度，而是取决于语主的文化心理的认同性。两种差别迥然的语言，由于文化心理的一致认同可以确定为一种语言；而两种差别微殊的语言，由于文化心理的拒绝认同则必须确定为不同语言。

语言的历史是语主的历史，语言的演变是语主的演变。从来不存在脱离语主的语言系统内部自身的自发演变，封闭社会的语言的变化几乎停滞说明了这一现象，而语主的变化可以说是一个耗散结构。一方面，由移民引起的语言历史分化，可以使方言演变成为亲属语言。另一方面，由移民引起的语主局部嬗变，也可以通过语言换用使所换用语言发生重大变化而成为一种方言。因此，语言演变是移民交融和语言接触的历史投影。

语言系统是人类文化系统中与其他子系统密切相关的的一个子系统。史前语言学既为史前，就不可能像现代语言学那样具有丰富的语料。以已有的各类语料为基础，采取语言学、考古学、历史学、遗传学、民族学、社会学的综合性研究方法，是探索语言史前状态的惟一途径。

（二）东亚南洋语言文化圈假说的史前构想

根据人类学和考古学的研究，人类在距今 4 万年前发生巨

大突变,从体质上的现代人逐步转变成为心智上的现代人。可以选择这一关键点,作为建构东亚南洋语言文化圈史前语言史的起点。距近 4 万年前,作为起源于新人的原始居民从中东出发,经过伊朗高原分别从南北两个方向进入东亚大陆。这些进入东亚大陆的新人,身材矮小且皮肤黝黑,和现代蒙古利亚人种不完全相同。北支从帕米尔高原向东进入黄河流域与大漠南北,然后继续向北,乃至经过白令海峡陆桥进入北美,逐步演化为东亚种群和北亚种群的新蒙古利亚人种,以及北极种群和印地安种群。其中东亚种群成为黄河流域居民的祖先,北亚种群成为大漠南北居民的祖先。

南支从印度次大陆[1],经由东南亚再进入南中国沿海和长江流域,成为蒙古利亚人南亚种群,继续向东南进入大洋洲的成为澳大利亚人种。其中南亚种群的北支——江淮类型成为长江流域居民的祖先,南亚种群的南支成为岭南—湄公河—恒河地区居民的祖先。如果汉藏语和印欧语确实存在发生学关系或一些远古同源词,除了可能源于最初欧亚人种的分化,就是分化后的文化扩散的借词。蒲立本认为印欧语和汉藏语的分化在6000 年前的说法之所以不能苟同,其理由就是人种体质类型在距今 3 万年左右形成,虽然不排除东亚现代人并非一次性迁徙,可能是一波波迁徙而来,但是没有遗传学证据表明白种人从欧洲来到东亚后再转变为黄种人。

越过白令海峡进入北美的印地安种群,成为说印地安语言的祖先。滞留在北极圈沿海附近的“海上游牧民族”的北极种

[1] 据研究,印度次大陆的土著是被从北方来的达罗毗荼人赶入偏僻从林的说南亚蒙达等语言的原始居民,后来达罗毗荼人又被从伊朗高原来的雅利安人向南方驱赶

群，成为说北亚细亚语的祖先，北亚原始语链的破裂而形成离散性语珠。东亚北部原始居民所说语言即原始北方语言。留居中国北方和蒙古、南西伯利亚一带的北亚先民，进入新石器时代以后，所说的语言逐步形成分布在大漠草原的原始阿尔泰语群，后来陆续向北方、西北、东北方向远徙。在黄河中上游的东亚人种所说的语言成为原始藏缅语群，后来分为三支：留在黄河中上游的北支成为氐羌语；进入缅甸的南支成为缅甸语；由川康进入西藏高原的西南支成为喜马拉雅语。东亚大陆南部原始居民所说的语言即原始南方语。进入新石器时期以后，不晚于距今7000年前，存在着北支江淮（南岛）语群和南支南亚（孟高棉）语群。南支南亚语群分布于恒河流域、湄公河流域、马六甲—尼科巴—苏门答腊以及云贵高原，后分化为蒙达语群、孟高棉语群和苗瑶语群。北支江淮语群，一部分沿江越淮迁到山东—辽东—朝鲜半岛，渡海到达日本列岛，在这一范围内形成东夷语群。一部分留居在长江以南的鄱阳湖—太湖流域，演变为百越语群（侗台语群）。一部分迁往南洋群岛甚至以远，演变为南岛语群（印尼语群）。

虽然中国原居民都是蒙古利亚人种，但是根据GM单倍型频率却可以分为南北两大类型，北方人群中GMag单倍型频率较高，在0.29～0.54之间，南方人群仅在0.03～0.25之间；与之相反，北方人群中GMafb单倍型频率较低，仅在0.08～0.39之间，南方人群却高达0.46～0.81。南北两大类型的地理分界线在北纬30度即长江一带[①]。与之一致，东亚大陆南北两支语言文化的显著区别是：作为基本词汇，北部叫水流为“河”，而南部叫水流为“江”；作为句法结构，北部语言为OV、AN的逆行

① 张余善编著《中国人口地理》，商务印书馆1997年。

结构,南部语言为VO、NA的顺行结构,但粘着结构却是其共同的原始特征。借助考古文化系统研究的成果,以中国境内三大考古文化系统的形成为座标,可以为东亚新石器时代的远古语言概貌的考察找到了一个相对支点。在三大考古文化系统的形成之前,南方是原始采集—狩猎演变而来的迁移型原始农耕经济,而后在一些地区又形成定居型发达农耕经济;北方是原始采集—狩猎演变而来的半农半牧型混合经济,而后在一些地区又形成专门的游牧经济。依据澳大利亚土著语言关系的性质可以推定,在东亚大陆早期农牧社会出现之前或原始采集—狩猎经济方式下的语言状况为连锁性原始语网,而农牧化引起的迁徙则是促使连锁性语网逐步分裂或重组的历史条件。南亚语、南岛语、藏缅语和阿尔泰语的部分底层相似性,也许正源于此种连锁性原始语网。

公元前5000年左右,三大考古文化系统的形成标志着青莲岗、仰韶、北方三大太古语区的形成,三大语区的共同特征是语言结构都是原始粘着型。由于中原地区得天独厚,多种语言文化的冲突交融促使原始华夏语的混成发生,因此包容着不同历史层次的孑遗。斯大罗丝丁之所以认为汉语和藏缅语关系最密切,其次是与北高加索—叶尼塞语,而和南岛语关系较远,就是因为依据的是氐羌化的周秦语料。沙加尔之所以认为汉语和南岛语同源,是因为利用的是搜集了许多南方词语的《集韵》,而《集韵》中搜集的用汉字书写的百越语或汉语东南方言词,正与南岛语的词具有对应关系。

白保罗认为台语与汉语词汇近似而与印尼语相差甚远,是上古汉语向澳泰语借词而后又贷词给澳泰语的结果,其前提——澳泰人的文化技术比北方华夏人发达——没有考古学证据。罗美珍的大陆台语先民换用原始汉藏语而残留着原始马来

语底层的假说和倪大白的台语“马来—支那语”假说，都隐含着台语是历史混合语的潜台词，而问题在于：台语先民是什么时间换用原始汉藏语？是什么条件引发了百越先民的语言换用或接受渗透？雅洪托夫提出，4世纪以后中原居民的大量南迁促使侗台语先民的语言转型或语言混合。作为中介将台语和印尼语联系起来的加岱语群，其中拉嘉语等在越南境内，黎语处于台语和南岛语之间，正好填补了台语和南岛语推移的中间环节。同样，关于苗瑶语是属于南亚语还是藏缅语的分歧，正好表明苗瑶语处于南亚语和藏缅语的中间性质。加岱语和苗瑶语的过渡性正是语言接触渗透的产物。雅洪托夫依据台语、加岱语与印尼语一致的词语在斯瓦迪斯百词表中只有15个，推定其亲属关系在台语和加岱语之间，甚至比汉语和藏缅语之间的关系都要远得多，以证明加岱语支并不存在。这一结论的不周之处是，既没有考虑语言渗透所引起的词衰减变量，更没有设想原始语网状态下的语言歧异因素。以单纯的语言分化论和封闭的语言年代学方法，也只能够得出这样的不尽符合语言历史的结论。

## 三、东亚南洋语言文化圈的变迁鸟瞰

距今7000年前，在中国境内大致形成的新石器时代三大考古文化体系格局，奠定了这以后中国境内乃至东亚—南洋语言文化发展的基本态势。三大考古文化系统的创造者都是蒙古利亚人种，所说的太古夷越语、太古氐羌语、太古胡狄语就是现代东亚—南洋语言文化圈内诸语言的祖先。

（一）现代阿尔泰语是太古胡狄语的后裔

在游牧文明和农耕文明的持久冲突和交融中，除了进入中原的阿尔泰人融入华夏汉族而换用了汉语以外，留居大漠南北的阿尔泰人，则保留着原始胡狄语的粘着结构、SOV和NA语

序以及语音和谐的特点。

粘着型的朝鲜语的系属比较模糊。处于朝鲜半岛的朝鲜语其基层为原始夷越语,基本词汇与南岛语同源[①]。从文化底层来说,朝鲜半岛—辽东半岛—山东半岛—江淮平原等,在远古同属水耕文化系统,所说语言是原始夷越语。从文化接触来说,一方面,战国后期的燕国和秦朝先后控制辽东半岛和朝鲜半岛北部,西汉初燕人卫满率领数千人进入朝鲜半岛北部,以燕、齐移民为基础建立了政权,扬雄《方言》收录了朝鲜语词,辽东—朝鲜半岛文化受到中原强势华夏汉文化的持续影响;另一方面,辽东—朝鲜半岛的北面是阿尔泰肃慎、乌桓、鲜卑等的游牧世界,从商周时代,原始朝鲜语可能就受到阿尔泰语的影响。由于文化接触,由此形成了以原始南岛语为底层,以阿尔泰语为表层,又向汉语大量借词的后代朝鲜语。在语言谱系研究中,以往的一些学者之所以主张将朝鲜语划入阿尔泰语系,其主要依据就是只看到朝鲜语和阿尔泰语的语言类型的一致性,而不了解朝鲜语的底层语言也是粘着型。朝鲜语即使归入阿尔泰语,也是形单影只,在N·鲍培的阿尔泰语关系图中与"楚瓦什—突厥—蒙古—通古斯—满洲统一体"[②]不相联系。因此朝鲜语是一种历史混合语,可以看作南岛语和阿尔泰语的中间类型。

(二)现代藏缅语是太古氐羌语的后裔

参照雅洪托夫的研究,原始华夏语、藏缅语和喜马拉雅语从太古氐羌语中分化出来的年代是公元前25或35世纪(炎黄时代)。卡伦语分化出来的时间稍晚,约在公元前20~25世纪之间(五帝时代)。藏缅语在公元前20到10世纪(夏商时代)开始

① 吴安其《论朝鲜语中的南岛语基本成分》,《民族语文》1994年第1期。

② 参见李增祥《突厥语概论》,28页,中央民族学院出版社1992年。

分化出各独立语支。缅语从彝语支中独立出来是在4世纪或6世纪(东晋南北朝时代)。

在原始氐羌语的子语中，有些(如嘉戎语)还有着丰富的粘着词缀，有些已经不多。现代藏语中粘着词缀还不少，动词的时态都用词缀表示。藏语还存在着被认为可能是后起的屈折方式，由此使藏语的语言类型表现为粘着→屈折→孤立的历史演变趋势。河野和桥本万太郎所说藏语在词汇方面和汉藏语大有关联，然而在句法方面更像阿尔泰语，如藏语的粘着结构、SOV型及NA、AN的并用等。词汇方面藏语和汉语大有关联，是因为源于氐羌的姜周使华夏汉语"氐羌化"；藏语的句法方面更像阿尔泰语，并非受阿尔泰语影响所致，而是原始氐羌语即粘着型而藏语保留着古老的句法结构。向远古追溯至距今一万年前，原始氐羌语应和原始阿尔泰语具有亲缘关系。

(三)现代南岛语是太古夷越语的后裔

南岛语系的分布区域相当辽阔，北起台湾，南到新西兰，东到南美洲以西的复活节岛，西到非洲东岸的马达加斯加岛。与世界上其他语系相比，南岛语系的数目较多，约有九百多种。中国台湾就有二十几种南岛语言。白保罗、奥德里古尔都推断南岛语源于中国大陆，从大陆出发南迁的居民将南岛诸语言流播到南太平洋群岛，马来—玻里尼西亚语和亚细亚—马来人的发源地应该到台湾和海南岛之间的中国大陆的华南去找。如果不是美国学者戴恩(Daien)以萨丕尔(1916)的"任何一群亲属语言的最早居民的故乡是现有多种语言的地区"的假设为前提，提出马来—玻里尼西亚语起源于美拉尼西亚的假说，那么"南岛语源于中国大陆"的观点应是公认的了。麦尔多克(Mekdok)认为，应当放弃在南岛语言和东南亚其他语系如台语、加岱语之间寻找亲属关系的尝试，中国南部——台语、加岱语的可能故乡——

距离美拉尼西亚太远[1]。麦尔多克根本没有注意到，南岛语东到复活节岛、西至马达加斯加的分布状况，表明南岛居民是非常善于海上迁徙的民族。而在从南太平洋到东非东岸的印度洋的广大区域中，中国南部正处于中间位置。

1984年，美国夏威夷大学白乐思(Robert Blust)根据语言分类法推论：南岛民族的故乡就在台湾。台湾语言学家李壬癸依据语言古生物学，进一步推阐南岛民族的扩散中心在台湾中南部[2]。除此之外，台湾南岛语扩散中心的这一假说的证据还有语言学原理和考古学证据。越早分裂的语言，彼此之间的差异越大；越晚分裂的语言，彼此之间的差异越小。越早分裂的语言，距离起源地越近，越晚分裂的语言，距离祖居地越远。其他南岛语区的纷歧性普遍不及台湾地区。南岛语区的其他考古文化遗址年代，都没有台湾发现的距今大约6000年前的大坌坑考古文化更古老。如果依据台湾考古文化推断南岛居民在6000年前就定居在此，那么6000年前的南岛居民的前身又从哪里来？这仍然需要用语言学方法，找出与南岛语言具有亲属关系的语言以及语主更早期的居住地域。比较可靠的证据显示，南岛语和南亚语很可能具有亲属关系，它们的共同起源地应是在亚洲东南部一带，台湾只不过是南岛居民迁徙的一个中转站。据白乐思推测，共同起源地大概在距今大约8000年以前的缅甸北部。凌纯声则认为南岛居民原居住在长江流域，尤其以左洞庭而右彭蠡为中心区域。最早见于中国古史者名为“九黎”、“三苗”，其后在西南曰“百濮”，在东南为“百越”。从考古学材料来看，古代中南半岛及南岛太平洋区域文化与中国南方的濮僚和

---

① 见伊万诺娃《论台语及其在发生学分类中的地位》。

② 李壬癸《台湾南岛语言的重要性》，台湾《汉学研究通讯》，1996年总58期。

百越民族具有密切的关系，与荆楚文化亦有关联[1]。

虽然，南岛语专家李壬癸不赞同沙加尔的汉语与南岛语同源说，但是，这一假说仍然得到了许多语言学家的首肯。因为汉南同源说不但有着欧洲语言学家探讨东方历史语言学的学术渊源，而且有着20世纪60年代以来人类学家、考古学家和文化史家对于太平洋岛屿文化和中国南方以及东南亚文化史前关系研究的学术背景。美国俄勒冈州大学人类学家杨江，根据语言文化、建筑风格、人种遗传、社团组织和宗教信仰的相似性以及文化习俗的共同性等方面的证据，提出越国人后代文化或"百越文化"与马来—玻里尼西亚文化有着远古时代的悠久联系[2]。早在6000年前，马来—玻里尼西亚人的祖先就开始了从中国东南沿海一带出发的长途迁移活动。途经台湾、海南岛、中南半岛，向南行进穿越菲律宾到达印尼。尔后一路向西，最远到达非洲东部的马达加斯加岛；另一路向东，到达南太平洋群岛和夏威夷、伊斯特岛。杨江甚至给出了一份东迁时间与到达地点的年代表：距今6000年，到达中国东南部和台湾；……距今4000年，到达密克罗尼西亚……距今1400年，到达夏威夷。除了部分人与当地土著巴布亚人混合之外，他们到达的都是荒无人烟的岛屿。一方面，他们上岛时带来了不少原先饲养的动物以及原产于东南亚的椰子、野芋等植物；另一方面，没有其他文化对之产生影响或相互冲突，因此保持着源于大陆东南部的文化习俗。

依据白保罗的加岱联盟假说，台语是留在大陆的一支远古南岛语的后裔。拉古勒里（T. Lakubeli）试图用台语的起源来解

---

① 凌纯声《中国边疆民族与环太平洋文化》，台湾联经出版事业公司1979年

② ［美］杨江《马来—玻利尼西亚与中国南方文化传统的关系》（吕凡译），《浙江学刊》1991年第1期。

释汉语和台语中的相似性，他把台语的形成说成是南部诸孟语言和北部诸语言——汉语和其他语言——的混合。按照他的研究，这一混合过程的结果就是台语词汇中包含了三分之一的汉语词汇，并且接受了汉语的单音节性和声调系统的特征。研究泰国中部语言的史莱格(Shlegel)也持有同样的意见。罗美珍和倪大白的意见与之相仿，只不过将语言混合改为语言换用，将底层语言诸孟语言改为马来语。苏俄莫列夫(L.H.Molev)提出，不排除汉语和台语起源于统一原始语言的可能性[①]。根据雅洪托夫的研究，汉语和台语的共同词汇数量相当多，但是见于斯瓦迪斯基本词表的却较少，修正表中完全没有，这些词汇都是较晚的汉语借词。雅洪托夫认为，可以用黎语作为佐证，它明显与台语具有亲属关系，但几乎完全没有来自汉语的借词，由此否认了台语通常被认为是汉语亲属语言的普遍看法。

粘着型的日本语系属一直比较模糊。在漫长的冰河时代，日本列岛和大陆相连。1949年，在群马县发现了岩宿旧石器文化遗址，表明日本列岛当时已有居民居住。分布在日本北方诸岛的阿伊努人，人类学家认为是蒙古人种南亚类型和赤道人种澳大利亚支系之间的混合类型。但阿伊努人也许并非分化后的南亚类型和赤道人种澳大利亚支系的后起混合类型，反而更可能是南亚类型和澳大利亚支系未分化前的原始太平洋人种或原始蒙古利亚人种祖先的后裔。

日本大和语，西方学者拉姆斯太特认为与高丽语和阿尔泰语近似，但松本广信在《日本语和南亚语比较词汇的研究》(1928)中认为，日本语与太平洋群岛的马来语和支那语关系密切。二者的分歧隐含着日本人是来源于北方的高丽人，还是来

① 见伊万诺娃《论台语及其在发生学分类中的地位》。

源于南方的马来人的推测。日本人源于高丽人还是马来人并没有本质的矛盾，因为高丽人与马来人都源于太古夷越人或原始南岛人。根据日本绳纹考古文化遗址都在北方，可以推断大和远祖是从北方即朝鲜半岛进入日本列岛的，然而他们并非是原始阿尔泰人，而是原先居住在大陆东南沿海的原始夷越人(南岛人北支)，他们的根在长江以南地区。日本早期居民的这一迁徙路线，和大陆水稻传入日本的路线(江淮—山东—辽东—朝鲜—日本)相一致。至于日本语和南亚语(实指南岛印尼语族)的相似性，是因为说这种语言的马来人也是从长江以南地区再南迁印尼—马来海岛的。大和的远祖，除了是从北路迁入日本列岛的原始夷越人，并不排除稍晚时期部分原始马来人从南路经台湾、琉球而后进入日本列岛。

除了新石器时代的早期移民，从夏商到周秦，中国东南沿海的居民仍然陆续迁入日本。倭人"男子无大小皆鲸面而文身，闻其旧语，自谓泰伯之后"(《魏略》)，其俗近江南吴越，又自称先周泰伯之后，当是在周代从江南由海路迁入。公元前三世纪，日本进入弥生文化时代，据一些学者考证，与秦人徐福东渡带来新的周秦文明密切相关。公元57年，东汉光武帝赐日本"汉倭奴国王"金印。五世纪前后日本使用汉字。中古时期，先后派遣隋使、遣唐使及大批留学生和僧人来大陆，学习中国文化。在汉字的基础上创制适合记录日本语言的假名，平安文化时代形成了民间假名文化和官方汉字文化的双层结构。长期以来，日本语虽然在语音结构和词汇系统方面深受到汉语的巨大影响，但是语言结构的粘着类型却没有改变。无论从原始大和居民来源还是语言类型，日本语当源于原始南岛语，为太古夷越语的一支后裔。也许正是列岛居民来源的南北交融性和历史层次性，才致使日本语具有与其他南岛语后裔不尽相同的特殊性，因此一些

日本学者主张日本语为混合语说。

(四)现代南亚语是太古夷越语的旁系后裔

根据白保罗的澳泰语系说,南亚语和南岛语皆源于原始南方语,分化的时间估计在新石器时期早期。现代南亚语的词音结构通常表现为一个主要音节,有时前面可再加一个次要音节。前缀和中缀比较普遍,但后缀不普遍。尽管现代南亚语表达语法意义的重要手段是词序,但可以推定原始南亚语是粘着结构。不可否认,南亚语从粘着型向孤立型过渡,华夏汉语的影响是最重要的因素。

作为古代苗蛮后裔的苗瑶族,秦汉魏晋居住在洞庭湖流域。南亚人最初的居住地在云贵以南,也许是在南岛人向东南迁徙以后,他们才从南部进入长江中游一带。随着永嘉之难以后中原居民的大批迁来,苗瑶族迁往湘西和云南。根据雅洪托夫的研究,苗语从瑶语中分化的时间是公元4世纪或10世纪,恰巧与历史上的永嘉之乱和黄巢起义所引发的中原居民南迁浪潮相吻合。在苗瑶语中,可以找到很多与孟高棉语共同的词汇以及少量的与台语共同的词汇。显然苗瑶语属于南亚语系,在语系内部最接近孟高棉语。苗瑶语中与汉语对应的词,多半是晚近即永嘉以后的借词。

系属比较模糊的越南语,和汉语的语言类型一样。雅洪托夫认为,越南语中有很多来自汉语的词,见于斯瓦迪斯基本词表的有13个,无疑这是较晚的汉语借词。越南语和台语关系复杂,越南语和台语相同的词有些是台语借词,但有些或与台语和孟高棉语,或与台语和印尼语分别具有对应关系。越南与台语的对应词见于斯瓦迪斯词表的有30多个,其中19个与孟高棉语或印尼语具有一致性。关于越南语的原语有两种假说:一是百越语说,一是南亚语(孟高棉语)说。前者如马伯乐的《越南语

历史语音学研究》(1912)，由于对越南语的单音节和声调有着如此深刻的印象，以致于他尽管面对越南语中占优势的孟高棉语基本词汇成分，还是假设了越南语和侗台语的发生学关系。后者如白保罗(1942)，他认为越南语作为孟高棉语群最东北的一个成员具有特殊的地位，越南语和孟高棉语系的关系正像台语和印尼语一样。由于汉语的影响，越南语和台语的词音结构大量简化、失去构形词缀，形成完整的声调系统。鉴于台语的词根在越南语里不是无足轻重的，这些变化必须归因于台语的影响。尽管如此，绝大多数的基本词根还是孟高棉语来源而不是台语来源。根据白保罗对台语的分析，可以类推出孟高棉语与越南语具有发生学上的关系。

越南北部居民，在秦汉典籍中称之为"南越"。秦兵的南征和建立象郡、南越王赵佗的统治、西汉的建立交趾、九真、日南三郡，都对越南的语言文化发生了深刻的影响。奥德里古尔所认为的汉语对台语和越南语的最大影响的公元 3－6 世纪[1]，正是永嘉之乱以后中原居民的大规模南迁时期。939 年，交州人驱逐南汉而吴权称王，越南从此成为独立国家，但是汉文化的影响并没有中断。这种持续的影响无疑是促使越南语，包括台语，类型孤立化的根本动因。根据雅洪托夫研究的结果，越南语同台语的对应词 30 多个，其中有 19 个与孟高棉语或印尼语具有一致性。也许可以这样解释：越南西南部的原语是南亚语，越南东北部的原语是百越语(或百越人南迁后所占据的原先南亚语区)，越南语是西南部南亚语和东北部百越语(台语)的历史混合语。

[1] 见伊万诺娃《论台语及其在发生学分类中的地位》。

## 四、华夏汉语和周边语言同源异流现象阐释

美国学者罗杰瑞和梅祖麟在《古代华南的南亚语》[1]一文中提出,一些不见于早期汉语甲金文记录的词语是源于原始南亚语或澳泰语的借词。通过汉语的"牙" * ŋra(上古)、* ŋa(中古)、ia(现代),与越南语"象牙"ŋa 和原始台语 * ŋa,以及原始茫语"獠牙" * ŋo'la 的比较,罗杰瑞认为这是来源于原始南亚语的借词。确定为汉语借自原始南亚语的理由是:1.最早的甲骨文献中记录的是"齿",其词义包含臼齿、动物獠牙、长牙和象牙。2."牙"比"齿"晚出。3.安阳殷墟曾发现象的遗骨,表明北中国有过大象,但在商周时象牙稀少且珍贵,主要是从长江中下游的越人地区进贡,如《诗经》:"憬彼淮夷,来献其琛,元龟象齿"。

这一结论颇有可以商榷之处。首先,依据某汉字不见于甲金文,就判断为当时语言中没有这个词的前提不能成立。假使如此,那么依据出土甲骨文所见 5000 多个字符,就会推出商人语言只有 5000 多个词语的荒唐结论。其次,根据《说文》"牙,象上下相错之形",其形如臼齿相磨之状;根据段注"统言之皆称齿、称牙,析言之则前当唇者称齿,后在辅者称牙",可见"牙"并非取象于象牙之形,"牙"指象牙是引申义。再次,因为不存在可与甲骨文比较的越南语等古代文献,因此无法证明是越南人等最先使用 * ŋa 这个词来专指象牙。根据泗洪的考古发现,距今 5 万年的淮河流域生活着菱齿象,根据殷墟发现的大象遗骨,表明中原地区也有过大象,即使商周时来自淮夷的象牙稀少且珍贵,即使原始南亚—南岛人可能最先使用 * ŋa 这个词来专指象

---

[1] Jerry Norman & Tsu-Lin Mei: *The Austro Asiatics in Ancient South China: Some Lexical Evidence*. Reprinted from Monument Serica, vol.32, 1976。

牙，也没有理由认为这就是借词，因为原始华夏族的族源包括古代黄淮一带的夷越人，由东向西推进的、导源于青莲冈考古文化系统的龙山文化就是原始夷越人创造的文化，殷商人就是西进中原的夷越人的一支，在原始华夏语中存在原始南亚—南岛人后裔语言中的词语并不奇怪。当然，如果仅仅从姜周时代的语言层面考察，可以认为 * ŋra 是殷商时代语词的承袭，但这一论断包涵着华夏语仅起源于藏缅语的预设。同样，在闻宥和张永言的文章中[①]，也列举出一些见于汉语书面语稍晚而存在于现代南岛—南亚语言中的对应词，并且将这些语词判定为上古汉语中早已完全"归化"的外来语成分。如"贝"，泰语 beer；占语、吉蔑语 bier；马来语 bia 等。商人文化源于东部沿海的海岱文化，"货贝而宝龟"，以贝为装饰和货币，没有理由认定"贝"类动物不是商人祖先原始夷越族的濒海生态环境中的常见生物，而一定是南海的舶来品。

依据东亚南洋语言文化圈假说和华夏汉语混成发生论，无论是侗台—苗瑶语，还是华夏汉语，其语言文化的动脉中都流动着太古夷越语的古老血液。如果这些词确实非常古老，也只能是太古夷越语的后裔所共有的，即是华夏汉语从"远祖"语言那里继承下来的固有词，而并非是从"弟兄"语言那里借来的外来词。之所以这些中外学者，将见于现在南方民族语而在汉语文献中"晚出"的词——不见于甲骨文而见于先秦时代并不晚——确定为上古汉语中的外来语成分或归化词，就是因为他们首先预设了建立在谱系树模式基础上的"汉语仅与藏缅语同源"或"殷商人也是藏缅

① 闻宥《语源丛考·鹘鸼鹭三词次第考》，《中华文史论丛》第4辑，上海古籍出版社，1980年。张永言《汉语外来词杂谈》，《语言教学与研究》1989年第2期；《语源探索三例》，《中国语言学报》1988年第3期。

人”，而与南岛—南亚语没有同源关系的前提。然而，如果立足于华夏汉语混成发生论，就可以改换一种新的思路，得出不同的结论。因此我一直认为，预设前提决定了材料的价值。

如果从语言史观追溯到更深层次的文化史观，之所以认定中国北部语言曾经向南亚语言借词，是因为首先预设了史前南亚文明可能曾高于黄河—长江流域文明。1966年以来，在泰国北部班清(Ban-Chiang)墓地考古文化的发掘过程中，尽管有一些美国考古学家通过测定死者骨骼周围的木片，推断套在人体手腕骨骼上的青铜手镯是距今6000年前的青铜器物，以致出现了青铜文明的发源地是西亚两河流域还是东南亚泰国北部的质疑。但是，后来对骨骼直接测定的年代结果却是在公元前300～50年之间，因此青铜手镯的年代，远远在中原商代青铜文明之后。尽管泰国北部盛产冶炼青铜的锡，但根据文化传播的理论，泰国北部后起的青铜文明可能源于更早的中原青铜文明，而不可能相反。

由此可见，依据汉语与南岛语发生学关系的证明，依据史前考古文化表明当时最高的夷越文明发源于江淮，并且溯黄河而上混合形成原始华夏文明共同体，而大陆居民向中国南部沿海以及南洋岛屿的迁徙在远古时期就开始了的考古文化学证据，“华夏汉语仅仅与藏缅语同源论”就不能够成立。因此白保罗在一系列文章中，试图采取南方语言文化向北方中原传播的思路

来建立新石器时期以来的大陆语言演变的模式的设想①，与东亚—南洋的文化史实和语言史实不相符合，马提索夫由此推定早期台语对汉语产生过巨大的影响的观点也同样不能成立。归根结底，为欧洲学术界至今仍引以自豪的历史比较语言学，其主要贡献——谱系树理论，只是欧洲语言学家根据印欧语料和印欧人古代经济生产方式所建立的一种语言文化演变模型。

① P.K. Benedict: *Thai, Kadai and Indonesian: A New Alignment in Southeastern Asia*, *American Anthropologist*（《台语、加岱语和印度尼西亚语：东南亚的一个新联盟》），1942；*Early Chinese Borrowings, Genetic Relationship, Diffusion, and Typological Similarities of East and Southeast Asian Languages*（《早期汉语借词》），1976　又 J. A. Matisoff 的观点见《马提索夫教授谈历史语言学和汉藏系语言的研究》，《美国语言学家谈历史语言学》（徐通锵整理），《语言学论丛》第13辑，商务印书馆1984年

# 第六章　语言的本质共性与语义语法学理论*

## 第一节　语法形态范畴的本质是语义认知范畴

### 一、人类语言符号与范畴化认知能力

语言是人类特有的认知符号系统、思维工具、交际工具和文化形态。以语言符号纽带，在神经生理和对象世界之间形成了人类精神世界，人们通过语言符号精心建构对象世界，对象世界也只有处于语言系统才具有价值。与后起文化不同，作为本原文化的语言具有本体和载体的双重属性，一方面作为认知符号系统的语言就是一种文化形态的本体，另一方面作为一切知识底座的语言又是其他文化形态的载体。人所认识的世界就是语言中的世界，人对语言的研究就是对反映对象世界的精神世界的研究，而这精神世界大致可以分为三个层面。

(一)心智机制层面

从体质的现代人类到心智的现代人类，经历了约10万年左

---

* 本专题研究是江苏省哲学社会科学"十五"规划基金课题"面向语言工程的语义语法学研究"的成果之一。

右。人类心智机制包括源于动物性反馈的心智感受力及其受控于神经生理的经验基础。认知语言学提出的所谓经验框架，就是植根于原始认知方式中的集体无意识。这些从距今几万年前的远古时代逐步形成的心智能力，一方面成为初民对"存在·感受"进行语言符号化的认知能力，另一方面语言符号产生后又反过来作用于心智能力，由此形成了语言符号和心智感受的互相塑造。

（二）语言符号层面

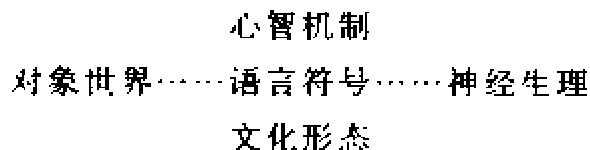
希腊哲人专注于讨论符号和符号所替代的事物是否具有同一性，因而产生了可知和不可知的迷茫。人类的可知是"不可知"中的可知，人类的不可知是"可知"中的不可知。从语言学家的立场出发，人类对客观世界的"可知"与"不可知"的界限，就在于是否存在以心智感受和经验框架为基础形成的语言符号，能用语言述说即已经实现符号化的对象就进入了可知，反之则否。正是依据这一语言符号系统，人类才从不可知的混沌进入了可知的清晰。除了可以借助语言符号来区分哲学上的可知和不可知，同样也可以借助语言符号来区分心理中的显意识和潜意识，凡是形成语义结构的意向就是显意识，反之则否。语言既是意识与对象之间的纽带，语言也是对象与意识之间的面纱。所谓"可知"，可以是"知多"也可是"知少"，可以是"知真"也可是"知伪"。语言既提供了人们认知对象世界的工具，也导致了对象世界与意识的隔膜。人类在不得不凭借语言的取景框来观察这样一个世界的同时，从而实现了历史上的第一次信息革命。

（三）文化形态层面

现代语言创造了现代人类和现代文化，现代语言的出现为人类社会的形成提供了必要条件，因此语言符号不但是人类与自然的纽带，而且是个人与个人、群体与群体之间的纽带。在文

字产生以前,人类借助语言符号,逐步形成了原始巫术和神化的史前意识形态基础;在文字产生以后,人类借助语言文字系统,在史前意识形态的基础上建筑起宗教、艺术、制度、道德和学术的意识形态大厦。联结对象世界和神经生理的语言符号思维有了一个文明社会的结晶,语言作为文化本体的同时又成为一切文化的载体。特别是在第二次信息革命的成果——文字出现以后,人类的文化形态以文字的物化记载而存在,对后世的文化发展具有不可摆脱的控制作用。从而呈现出心智机制、语言符号和文化形态互相作用的动态历史过程。

无论是在对象世界和神经生理之间,还是从心智机制到文化形态,语言符号始终处于中介位置并起着纽带作用,图示如下:




公元前的几个世纪,希腊哲人在探求客观世界和主体认知的关系中,萌发了古希腊语法研究,从哲学角度对语言意义的思考成为语言研究的原初对象。在这一探索中,当发现形态变化与实词类别具有印证性、实词词类与句法成分具有对应性以后,希腊语法学便逐步买椟还珠,致使西方语法研究走向了形态语法学的道路。与并非完全排斥语义的传统形态语法学相比,从形式分析语言的结构主义或美国描写主义则完全摈弃了语义,而乔姆斯基的生成语法学则把句法结构规则当作基础部分以研究人的语言能力。以研究对象为标准,西方语法学可以分为形态语法学、结构语法学和形式语法学三个阶段,两千多年来的注重语言形式的西方语法研究,一言以蔽之:语形语法学。

语形语法之所以可以作为语言理解和使用的工具,这是因

为人际交流中，已经存在一个建立在经验基础之上的不很清晰但便于检索的庞大语义网络。但是，无论是描写静态的结构分析，还是从事转换的语言生成，由于摈弃或轻视语义结构网络，结构分析则徒有其表，语句生成实言不由衷，以致于进行人机对话时，由于计算机不存在人类认知的“先天”语义网络，因此依据语形语法就无法自动分析和正确生成符合人类感知的语言。自然语言的计算机信息处理研究，向两千多年来、迄今为止仍然占据主流位置的西方语法学或语形语法学提出了根本性的挑战。

由此看见，西方语言学传统一直以为反映语言结构系统的语形规则，不过是语义结构在句法表达层面的投影。人类语言的共同本质属性是语义性，语言结构的本体是语义结构体，语言结构的真正研究对象是语义结构单位和语义结构关系。一旦认识到这一事实，自然会重新认识西方传统语法学赖以建立的主要基础和重点研究对象——以形态变化为显著特征的语法范畴(grammatical category)的本质。传统的语法范畴或词法范畴，通常被定义为“词的形态变化所表现出来的语法意义类型”，而词的形态变化则被认为是语法形式。语法范畴并非像通常所认为的那样表现语法意义，形态变化并非像通常所认为的那样作为语法形式。如果能理解，人类的范畴化能力是建立在经验基础之上的认识对象世界的基本手段，就会明白——语法范畴的本质是先民通过一定语音要素表现出来的原始语义范畴，简而言之，语法范畴的原型就是原始认知范畴，或以语音要素为标记的原始语义类别。

作为人类对世界认知的概括表现或语言演变中的语法化的重要方式，语义的形态范畴化在人类所有语言中都经历过，并且绝大多数语言将先民的这份原始认知遗产作为语法范畴一直承传下来。只有充分揭示语法形态范畴的认知范畴和语义类别的

本质,才能从根本上动摇语形语法学的基石,重建符合人类语言本质的语义语法学大厦。

## 二、事物的原始认知范畴或语义类别

名词(含代名词)的形态范畴反映了先民从事物的分类、计数和时间等,到立足于事物以观察事物与动作之间的关系或事理关系的原始认知过程。

(一)自然性别的性范畴

虽然性(gender)范畴不可能是一个现代生物学或科学生物学概念,然而是一个通过范畴化形式反映出来的原始生物学概念,很多语言中的性范畴与动物自然性别大体一致,可以证明这一点。拉丁语、现代俄语和德语是阴、阳、中三分法,现代法语是阴、阳二分法。在拉丁语中,阴、阳基本上是动物的自然性别的分类,中性或共性是既可作阴性也可作阳性的名词,如公民、婴儿、双亲、伴侣和领袖等。以现代科学眼光来看,许多语言中都存在性范畴与动物自然性别的龃龉之处。拉丁语中的阴性或阳性名词中都包括一些无生命事物,如河流、山脉的名称属于阳性,而国家、城市的名称却属于阴性。然而,这些现代科学中无生命的事物,可能在先民的心目中是“有生命”的事物,也就是说与最初的符号隐喻或万物有灵论有关。“罗马”在拉丁语中为阴性,则与“罗马城”的起源有关,“罗马城”的神是“母狼”。由此将“罗马城”认定为阴性,并且类推到其他城市。现代德语的妇女(das weib)、少女(das madchen)和古英语的妻子(wif)和妇女(wifmann)是中性。这些令人难以理解的性范畴,都应当从原始神话、民间习俗和词语演变等方面去寻找理据,共时语言学只会蒙蔽人们。

面对诸多现代语言,可以发现含义相当的同一词语在不同

语言中却划入不同的性范畴：

| | 现代法语 | 现代德语 | 现代俄语 |
|---|---|---|---|
| 太阳 | le soleil | die sonne | солнце |
| | 阳　性 | 阴　性 | 中　性 |

这些相互龃龉现象，也许反映了不同语主的认知差异性或语言流变的歧异性。

（二）事物数量的数范畴

具体事物具有数（number）范畴，而原料事物、集合事物和抽象事物没有形成数范畴，这一状况反映了先民的数量认识从可数事物开始。现存不同语言的数范畴种数并不一样。虽然在许多语言中存在单数和复数的二分制，但现代爱斯基摩语、藏语和景颇语却存在单数、双数和多数的三分制。古印欧语的数范畴是三分制，而现代英语却从古英语的三分制蜕变为单复二分制，可以推定三分制是古印欧民族的原始计数形式。南岛语系的美拉尼西亚的安纳图（Annaton），一些语言还采取有单、双、三、多数的四分制。语法范畴保留着历史上某一阶段曾通行的原始计数制度。

根据汉字独木为“木”、双木为“林”、三木为“森”，以及“雥”（三只鸟）、“粗”（麤，三只鹿）、“晶”（三颗星）等会意字，“道生一，一生二，二生三，三生万物”（《老子·四十二章》），以“三”表示多也许是三分制原始计数方式的孑遗[①]，可以旁推华夏汉语可能经历过“数”范畴阶段。人类的计数认知从一开始，到二，再到三或多，或者进一步发展到五进制、十进制或十六进制等等，三分制可能是一个关键。通过数范畴的歧异，可以看出先民原始计

① 参见杜勤《“三”的文化符号论》，29页，国际文化出版公司1999年。

数认知发展的不同阶段。

(三)事物存在的时范畴

一些语言中存在表现名词时制的时(tense)范畴,以表示同一事物不同时间内的存在性质,如美洲胡帕(Hupa)语、爱斯基摩语和非洲豪萨(Hausa)语:[①]

| | | 过去 | 现在 | 将来 | 完成 |
|---|---|---|---|---|---|
| 胡帕: | 妻子 | xouɬ*neen* | xouɬ | xouɬ*te* | |
| 爱斯基摩: | 烟 | puyo*thlu*k | puyok | puyo*qka*k | |
| 豪萨: | 我 | na | mā | | nā |

即使是同一事物,在不同的时间段则可能具有不同的性状。哲学上的名言"一个人不可能先后跳进同一条河流",反映了同一条河流的不同的时间段,同样,依据名词时制而强调某一事物的差异性,也可以进一步说:"同一个人不可能先后跳进同一条河流"。

(四)事物特征的"类"范畴

在世界语言中还广泛存在着不同特征的多种类(kind)范畴或类别范畴,其中以形象特征认知的类别范畴最能反映原始认知的意向性。同时,在一些语言中还存在一些包括数范畴的复合性或多层级类别范畴。表现物体自然性别认知的性范畴,其实也属于类范畴之一,不过是欧洲语言比较普遍且传承下来的类范畴。

**1.有灵与无灵的类范畴**

西班牙语有灵类前加范畴标志,无灵类前不加范畴标志:

---

① 语法范畴的用例,转引自彭楚南《语法范畴》、王振昆等编《语言学资料选编》(上册),中央广播电视大学出版社。

母亲爱女儿 La madre ama a la hija.(hija"女儿",a为有灵标志)

母亲爱花 La madre ama la flor.(flor"花",零形式为无灵标志)

阿巴扎(Abaza)语名词分为社会—主动(有理智)和社会—被动(无理智)类,这一划分也投影在数词以及其他词类中。

2. 高加索诸语的类范畴

高加索诸语言的类范畴从2~8个不等。拉克语(Лакский)具有五种类别范畴,与动词的一致关系表现在系词"是"上面、形容词、数词和代词也相应地分为五类:

| 类范畴 | 单数形式 | 复数形式 |
|---|---|---|
| 男人类 | Ур | бур |
| 女人类 | дур | бур |
| 混合类 | бур | бур |
| 东西和手制品类 | дур | дур |
| 屋子 | бур | дур |

其中女人类,以前包括女人的名称,现在只剩下少数几个名词,其他的都转入了混合类(可能由于形态的变化或范畴标志的混同,并非由于类别性质的改变)。混合类原来只包括动物、非生物,也许动物和非生物在语主的原始认知中原为一类;也许非生物原来单独一类,由于范畴标志的变化而与动物类混同一类。依据复数形式的相同可以推断,男人类(包括拟男物)、女人类、原动物类属于有生命类。依据女人类和手制品单数形式的相同,可以推断手制品原来是女人的产品。"屋子"(къатта)单独一类,可能其建筑材料来自非生物,因此单数形式与混合类相同;可能其建筑过程与手工制作类似,因此复数形式与手制品类相同。

3. 班图诸语的类范畴

非洲班图诸语言的类别范畴从十几种到二十三种不等,非

常丰富。鲁干达(Luganda)语的类范畴分单复数,用不同的前缀来表示:

| 类别范畴 | 单数形式 | 复数形式 |
|---|---|---|
| 生物人类 | omu- | aba- |
| 生物植物 | omu | emi- |
| 生物动物 | en- | en- |
| 无生物 | eki- | ebi- |
| 石头,船等 | eri- | ama- |
| 对称的(耳、腿) | okv- | ama- |
| 长而薄 | olu- | en- |
| 高而瘦 | olu- | ezi- |
| 巨大类 | oyv- | en- |
| 细小类 | aka- | obu- |

人类和植物的单数标志相同,可以推断鲁干达祖先生活在植被世界,与植物关系密切,而与动物相对区分。其中动物类包括手制品"针"和"鼓",可能因为制作针、鼓的原料分别是动物的骨头和皮子。"石头"和"船"等,可能概括为工具类。长而薄类包括"头发、桥、街道和日子","头发"的长是具体物体的长,而"日子"的长则是经验感受的长。细小类除了细小的东西,还包括看上去不快的东西和人的不愉快感,以致于"危险"也归入其内,也许这一复杂的类范畴来自于原始生活环境中某一具体细小东西对人的经常性伤害(刺伤),由此形成的"细小的——不快的——危险的"的心理反感积淀。这些都反映了语主意识形态中的特殊联想或符号隐喻。

曼得(Mande)语有可分离(alienable)和非分离(inalienable)

的类范畴，可分离类物体有财产、房屋、儿女、妻子；非分离类物体有身体的各部分、父母、丈夫。如：

（我的）妻子　*na* musu　（na 是可分离标志）

（我的）丈夫　*n* kai　（n 是非分离标志）

彭楚南将 alienable 和 inalienable 翻译为"可转让"和"非转让"[①]，似不贴切。从个人的角度对相关事物进行分类，财产和房屋是身外之物，但身体的各部分却不可分离，也许这是"分离"范畴的原始出发点。父母儿女，儿女离不开父母，但父母可以让儿女离开；夫妻之间，妻子离不开丈夫，但丈夫可以让妻子离开，反映了原始"父权"或"夫权"的意识形态，表明这一类范畴出现时男子已经成为家庭的主导。类范畴是非洲语言中最普遍的语法范畴之一，这些语言中的一些"离奇"范畴，恰恰反映了这些语言的最初语主的原始认知特点和意识形态。

**4. 吉俩克语的类范畴**

库页岛的吉俩克（Giliak）语的类范畴，表现在数词和名词的一致上，以数词"一（词根 n-）"为例，至少有下列这些变化（括号内为相当的汉译）：

| 类别范畴 | 名　词 | 词形变化 |
|---|---|---|
| 对称类（一双） | 眼睛、耳朵、手、脚、足印、鞋子等 | nv*asq* |
| 生物类（一只等） | 蝇、昆虫、鸟、鱼、动物等 | nən |
| 长条类（一棵等） | 树、木棍、草、植物等；路、小河等 | n*lex* |
| 薄平类（一张等） | 皮子、席子、烟叶子、纸等 | n*rax* |
| 小圆类（一颗等） | 牙齿、鸟蛋、拳头、斧头、星星等 | n*ik* |
| 大圆类（一个等） | 太阳、月亮、湖泊、大河、房屋等 | n*aqr* |

这些类别的范畴化程度已经较高。在吉俩克语中，与吉俩克人

① 彭楚南《语法范畴》，王振昆等编《语言学资料选编》（上册），468 页。

及其生产方式最密切的每类事物都有一个类别标志：

人神类：(一位)nin

捕鱼类：木板(一块)net；船(一只)nim；捕鱼具(一套)nfat；

渔网(一张)nvor；网目(一个)niu；网线(一根)nerqe；

鲇鱼(一串)nqaq；鱼干(一束)nar

狩猎类：雪橇(一张)nir；地方(一块)navr；绳织物(一些)nlaj；网(一张)new

度量类：一庹 na；一拃 nma；熊脂厚度(一层)nit

这些可能更加古老的原始类范畴，不但表明了吉俩克人的原始信仰，同时还描绘了库页岛上原始的渔猎生活场景。以捕鱼和狩猎为生的吉俩克人，他们信仰的神是人格化的事物，因此与人属于一类。有时他们带着一套捕鱼工具，划着木板制作的一只小船，用渔网去捕鲇鱼。回来以后，一方面把一串鲇鱼晒成一束鱼干，另一方面用一根网线修补好渔网上损坏的一个网目。有时他们乘着雪橇到一块地方去，用一些绳织物编结的一张网捕熊，杀死熊以后取出一层熊脂。吉俩克人测量物体长度，较小的物体就是手指张开的一拃，较大的物体就是两手臂伸直的一庹。

吉俩克人对事物形状的范畴化表现，以及数词依据名词的类别进行一致性形态变化的特点，对研究汉语量词的形成具有启发性。汉语的"量词"其实不表量，事物的数量由数词来表示。汉语的名量所表现的是名词所指称的事物的形象类别的特征，因此应当叫做类别词或"形类词"。汉语量词的形成实际上包括了两个问题，一个是数量名结构的来源，一个是单个量词的语源。在吉俩克语的类范畴中发现了数量名结构来源的线索，数量名结构的形成导源于对事物的形象分类，同时也反映着从个别事物的形象认知扩展为原始认知同类事物的形象认知。吉俩克语通过附着在数词后的形态变化所表示的形象语义类别，在

汉语中通过孤立化的量词来体现，至于量词的出现是否作为数词形态变化空位的补偿，因为汉字不是表音文字，上古数词是否曾经具有屈折变化无考。

（五）事动关系的格范畴

与性、数、类等范畴所反映的对事物的原始归类不同，格(case)范畴则是通过形态变化反映在符号配列中的，从事物立场所观察到的事物与动作之间所具有的关系上的原始认知烙印。名词格范畴中的领属、施事、受事、与事、地点、时间、工具、原因、比较、称呼等的分类，反映了先民对物体的领属关系，物体由动作所支配的施、受、与关系，时空与动作的存在关系，工具与动作的支配关系，原因与动作的因果关系，不同物体的比较关系，执行者与被执行者关系的认识。格范畴是最复杂的语法范畴之一，不同的语言可能具有不同的格位，人类语言的不同格位高达一百几十个。彭楚南《语法范畴》中列出了 22 种格，如主格、属格、与格、对格、造格、前置格、呼格、夺格、作为格、部分格、成为格、暂时格、缺乏格、伴随格、方式格、直到格、相关格、原因目的格、时间格、主动格、比较格和一系列的方位格。

现今不同语言的格范畴种数差别极大。闪含语系的阿拉伯语有 3 个格。印欧语系的拉丁语、俄语、德语分别存在 7 个、6 个或 4 个格，而英语的人称代词仅保留了 3 个格，名词只剩下通格和所有格，现代英语也可能是欧洲语中格范畴最少的语言。阿尔泰语系的维吾尔语 6 个格，蒙古语 7 个格。藏缅语中的藏语 8 个格。属于乌拉尔语的芬兰语和匈牙利语分别多达 16 个和 25 个格，而高加索语的格鲁吉亚语则现存 23 个格。如果在历史上存在着格范畴由多变少的趋势，也许这些格范畴种数多的语言，其语言年代或结构类型较为古老。

## 三、事件的原始认知范畴或语义类别

名词格范畴反映的是从事物立场观察事物与动作之间的关系,而动词的许多范畴则反映了先民立足于动作事件,对与动作事件相关的事件角色、事件时间、事件状态、事件施受和事件虚实的范畴性认知,并通过语义符号配列中的形态变化表现出来。

(一)事件角色的称范畴

人称(person)范畴既反映了述说者对事件角色的己方、对方或他方的区别,又反映了述说者对事件角色的数量的分类。现代英语代词有三种人称,其中仅仅剩下第三人称单数要求动词进行相应的形态变化,而现代维吾尔语仍然有完整的动词人称范畴。

(二)事件存在的时范畴

与名词的时(tense)范畴表示事物的不同时段的属性不同,动词的时范畴反映了以现在为坐标观察事件存在的时间前后,以确定所述事件和述说者之间的时间关系的认知。拉丁语有现在、过去和将来时三种时范畴,反映了欧洲人以现在为坐标而区分已然和未然的双向延伸时间观。

(三)事件执行的体范畴

体(aspect)范畴反映了事件的持续进行还是短暂进行状态、完成还是未成状态、一次动作还是多次动作等,来确定所述事件的执行状态的原始认知。常见的体范畴往往是配对的,如:经常体——暂时体;完成体——未成体;进行体——未行体;一回体——多回体;反复体——非复体,另有一些是与普通体相对的,如:随便体、迅速体、加强体、结果体等。现代英语有一般、完成和进行三种体,现代俄语的有完成体和未成体、一回体和多回体、定向体和非定向体,但现代法语和德语却没有体范畴。由此

# 第六章　语言的本质

| 三分制 | 近指 | 中指 | 远指 |
| --- | --- | --- | --- |
| 日本语 | kore | sore | are |

| 四分制 | 这(近) | 这(远) | 那(近) | 那(远) |
| --- | --- | --- | --- | --- |
| 艾伊努语 | ta-an | nei-a | nei-an | to-an |

有些语言的距离性空间定指,还附加方向区别。如:

| | 这一个 | 北边那一个 | 南边那一个 | 下边那一个 |
| --- | --- | --- | --- | --- |
| 爱斯基摩语 | manna | anna | qanan | kanna |

蒙达语族的散塔里语的指示代词分为四级:1.最远、近、远;2.有生物、无生物;3.单、双、多;4.视觉性、听觉性。指示代词总共有两三百个之多,可以表达最细微的复合性空间定指和多种差别。

(二)动静性空间区别

事物的静止和动作的运动和相对,一些语言中存在着与动作密切相关的空间或方位的形态范畴,通常区别为空间静止式和空间运动式。空间静止式表示"在何处"、"在这里",空间运动式表示"向何处"、"向这里"。如:

| | 静止式 | 运动式 |
| --- | --- | --- |
| 彝语"何处" | xa-lo | xa-o |
| 俄语"何处" | где(г<к) | куда |
| "这里" | здесь(з<с) | сюда |

(三)细微性空间区别

阿留申语(Aleut)中的方位认知范畴最为复杂。以"我"为参照,"他"具有几十个方位变化:

第一个他 第二个他 第三个他 更远的他 倒数第二个他

wan ŋan ikan akan qagan

坐在我前面的他 我正对面的他 在上面的他 在正上面的他

qikan qakan ikan aran

在下面的他 在更下面的他 在最下面的他

unkuan unan sakan

复杂的方位范畴反映了这些语主对空间和方位的原始认知。这些十分细密的原始空间认知,一方面,反映了当时还没有产生距离的计量尺度,因此以参照物来确定目标物的相对距离感;另一方面,空间和方位的认知复杂性,反映了原始认知范畴化能力的具体性和个别性。

属于高加索语言的拉克(Laks)语的方位格有两重性,前置类的"运动方式范畴"按照意义可分为"在、到、向、通过、由"5类,后置类的"空间位置范畴"按照意义可分为"里面、上面、下面、后面、旁边、附近"6类,两相复合可以得出30个方位格。通过拉克人对事物存在方位的细微区别,可以探讨拉克人先民的空间范畴化认知能力。

(四)级差性状态区别

状态形容词和状态副词的级差性区别认知形成级(degree)范畴。状态形容词的级范畴是先民对同类事物的状态,进行程度比较所形成的级差性语义认知范畴。现代英语有原级、比较级和最高级三种,现代蒙古语形容词有原级、弱程度、强程度、比较级和最高级的五种。在欧洲的一些语言中,形容词为了在语法上和名词保持一致,也沾染了名词所表达事物的性范畴。这反映了在先民的认知中,具有不同性范畴的事物也应具有不同性范畴的性质。状态副词的级范畴是反映先民对同类动作的状态进行比较所形成的级差性语义认知范畴。许多现代语言保留

着原级、比较级和最高级的区别,但是在智利的西班牙语中,状态级差却多到12级。

## 五、认知语义范畴化的基本特点

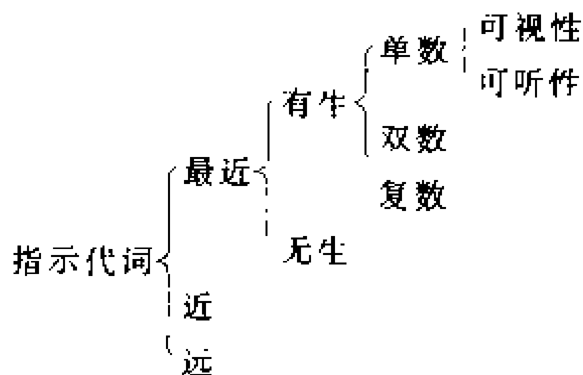
(一)认知语义范畴的层级性

通过对社会发展和人类认知的历史考察,可以推定,传统语法范畴或语义范畴是新石器时期人类的原始认知系统,通过这些范畴窥见的是原始认知过程和意识形态系统。人类对客观对象的感知可以大致概括为事物类认知和事件类认知。对具体事物的抽象形成个别名词,对具体事件的抽象形成个别动词,对事物的性质抽象形成形容词,对事件的状态抽象形成状态词。

从纵向考察,人类对客观对象的认知是一个不断抽象的过程。个别事物的命名过程是个别事物的符号化,形成个别事物词或个别事件词;特征相似和日常相关导致符号的联想化,个别词推衍孳生出一系列新词,即在联想基础上的再抽象或概括;在新的抽象基础上,对具有相同性的事物和事件、性质和状态的进一步概括或范畴化,则形成认知范畴。符号化、联想化和范畴化能力推动着人类认知的不断深化,由此形成人类认知的三级抽象模式,在认知网络中表现为符号化、联想化与范畴化认知层级,在语言网络中表现为个别名称、属性标志和语义范畴层级。从横向考察,人类对客观对象的认知具有相关发散性或延展性。以事物为核心的认知,向具体环境的发散延展概括出性别、数量、类别、时段定指、空间定指等属性或范畴。以事件为核心的认知,向具体环境的发散延展概括出时间、空间、状态、施受、虚实等属性或范畴。同类的事物合伙同类的事件但具有比较的可能,状态差别性认知分别形成形容词和状态词的级范畴。

这些认知范畴化能力和语义符号化能力,在语言中形成一

个个语义类别或语义范畴，在层级体系化能力的基础上，建立起控制着句法结构的原始认知语义结构网络。这种认知语义结构网络，具有层级性和网络化特征，如蒙达语系的散塔里语的"指代系统"：




通过层层叠加性组合，由此形成两三百个表达细微的不同语义差别的指示代词。然而，同时带来了语义范畴和形态变化的复杂性。根据彭楚南的计算，在理论上，一个动词如果有六种范畴，那么就可能有：3(时)×2(完成)×2(进行)×2(经常)×3(人称)×2(数)×3(性别)×2(态)×4(式)＝3456个变化形式。如果语法形式像粘着词缀一样具有规则性，那么只要掌握3(时)＋2(完成)＋2(进行)＋2(经常)＋3(人称)＋2(数)＋3(性别)＋2(态)＋4(式)＝23个语法形式，就可以驾驭这3456个的可能变化[1]。但是屈折语的实际情况并非如此。

（二）认知形态手段的像似性

认知语义范畴必须借助一定的语音方式即形态手段来表现。从现存语言来看，闪含语、语言年代古老些的印欧语、澳大利亚土著语、北极土著语普遍使用交替手段，阿尔泰语、乌拉尔

---

① 彭楚南《语法范畴》，王振昆等编《语言学资料选编》(上册)，502页。

语普遍使用附加手段,南岛—南亚语普遍使用重叠手段。不仅语义范畴具有原始认知性,而且表现这些语义范畴的形态手段,也与所反映的世界图式之间具有某种结构像似性。

1 表示增多的重叠和附加

音节重叠可以增加词长,与许多物体的存在呈现数量增多之间具有结构像似性。因此表示数量多的名词复数用单数原形的重叠来表现。如:

| | | 单　数 | 复　数 |
|---|---|---|---|
| 藏　语 | "谁" | su | susu |
| 马来语 | "人" | oraŋ | oraŋoraŋ |

同样,在一些语言中,表示程度加深的形容词或状态词的级范畴用原级的重叠来表示。在用附加手段表示名词复数的情况下,这些语言中的复数词形总比单数长。用附加手段来表示形容词的级,比较级和最高级的词形总比原级长,最高级的词形总比比较级长。

2. 表示变化的交替

不变表示事物或事件的稳定性,而交替表示关涉事物或事件的差异性。通过某一词音结构中的部分语音要素的交替,可以表示同一事物或同一事件的些微变化。古希伯莱语的词干一般由三个辅音构成,通过不同元音的交替表示不同的语法意义。辅音词干的固定不变,以表示这些相关事物或事件的核心语义的一致性,而辅音之间的元音交替,则显示了这些相关事物或事件的部分语义的差异性。如:

k-t-l　辅音词干

kotel　"杀"的施动者——杀人者

katul　"杀"的受动者——被杀者

ketol “杀”的不定式

满语的性范畴，通过语音交替来表示。如：

| | 阳性 | | 阴性 |
|---|---|---|---|
| 父亲 | ama | 母亲 | eme |
| 男人 | haha | 女人 | hehe |

父亲和母亲的共同点或共同义位用语音-m-表示，不同点或区别义位用语音 a-a 、e-e 区分；男人和女人的共同义位用语音 h-h-表示，区别义位用语音 a-a 和 e-e 区分。在名词数和形容词的级范畴中，也存在交替手段。虽然在词长上没有变化，但仍然反映了词干的不变性与事物的同一性之间的结构像似性，反映了元音交替的可变性与事物数量或程度增减性之间的结构像似性。

（三）认知语义范畴的消长性

原始认知语义范畴刻下了先民认知对象世界的一步又一步的足迹。然而，认知语义范畴的原始性、细密性和复杂性，注定了这些范畴必然经历弱化、混化、虚化而数量逐步减少的总体演变趋势。根据性范畴中隐含的万物有灵论和数范畴中的简单计数法，可以认为语义范畴的形成是从事物或名词的性和数范畴开始的，经由名词的格范畴，一方面扩展到表示事件的动词的诸范畴的发展，另一方面扩展到修饰事物的形容词和描绘动作的状态词的级范畴。同样，语法范畴的消失可能也是从性范畴开始，再波及到数和格范畴。现代英语名词的性范畴已经基本消失，除了少量表示人或动物的名词，如 actor（男演员）—act*ress*（女演员）；tiger（公老虎）—tig*ress*（母老虎），仍然残留着古英语的阴性词尾-ess。数范畴从古英语的三种变为现代英语的两种，并且复数形式有可能在几十年内消失殆尽。格范畴从古英语的四种只剩下现代英语的所有格一种，而且所有格的词尾 -'s 早

已丧失刚性规约,一些名词的所有关系早用分析形式 of 来表示。

根据不同语言中现存范畴的多少,可以认为认知语义范畴有一个从少到多,再从多到少的彼此消长的过程。在原始认知语义范畴萌发阶段,应是先从无到有、后从少到多,与先民的范畴化能力萌发、由此及彼相平行;而在原始范畴系统形成以后,随着抽象能力的提高或认知体系的转型,一些早期的原始认知范畴可能逐步混淆或磨灭,在某一阶段以后,又转向范畴种类的从多到少,形态变化的从强趋弱。根据现存语言中格范畴的差异较大,可以推断其在近 3000 年来存在着从多到少的演变趋势。除了认知能力的发展和知识系统转型,还要考虑到语言接触融合的因素。不同形态范畴系统的反复撞击只有同归于尽,因此融合程度越高的语言则范畴失落越多。而那些原始认知语义范畴复杂的语言,都是封闭的语言,其语言年代应较为古老。由于语言发展的不平衡性,即使是同一祖语分化而来的亲属子语,结果几千年的各自发展,其形态范畴的演变也各不相同。英语和德语同属东日尔曼语,现代英语名词 2 个格,残留人称代词 3 个格,而现代德语是 4 个格。属于斯拉夫语族的俄语 6 个格,而属于罗曼语"死语言"的拉丁语则 7 个格。

(四)认知语义范畴的虚化性

语主认知能力的发展和知识系统转型的历时变化,可能导致原始语义范畴出现两种结果:如果古今认知观念一致,某些原始语义范畴今人就容易接受;如果古今认知观念不一致,某些原始语义范畴今人就无法理解。一些语言学家对名词"性"范畴大惑不解,在以现代生物学体系去批评"性"范畴的"自然性别混乱"时,其前提可能认定性范畴从产生之初就只是纯粹的形态,而没有考虑到先民的性范畴与现代生物学属于不同的认知体

系。通常认为“代表词汇意义的是‘词根’，代表语法意义的是‘形态’”，其实从原始认知立场，形态所表现的语法意义同样也是词汇意义，或原始认知语义的必要组成部分。在保留着性范畴的语言中，如“父亲”和“母亲”、“男人”和“女人”等词语，其中的性范畴始终保持着词汇意义，而在另外一些词语中，曾经作为原始语义的性范畴才不再与后世的词汇意义一致，因此被看作与词汇意义无关的抽象化“语法意义”。然而，这些词的类别特征并没有完全消失，分布位置与结构功能的一致性，仍然把这些具有同一范畴的词汇聚合在一起。面对这一参差不齐的情况，以致于在后人眼里，原始语义范畴仅仅是一些形态的空壳，也许就是所谓语法化(grammaticalization)或虚化(bleaching)的结果。可以认为，所谓语法化不过是原始语义范畴的语义失落，或原始知识系统的后世转型。

一方面，无论是古今认知观念的一致还是龃龉，都可以断言——传承到现代的形态语法范畴所表现的是历史上的语义。之所以希腊语研究中形态范畴成为词法分析的形式或词类划分的标准，就是因为形态变化是语义结构的句法表层投影。另一方面，由于语音要素的历时变化可能导致标记的错位、混同以致失落，从而造成某些认知语义范畴的交错、合并以致消失。虽然原始认知语义范畴逐步语法化为后世的语形范畴，但仍然可以看出其中保留的原始认知烙印。人类语言的本质是语义性而不是语形性，貌似形式标记的语法范畴可以明确地定义为历史上的语义范畴。

## 六、语法形态范畴的消亡和继起的语法化

一种语言的语法化类型可能是历史上语法化类型的传承，也可能是旧的语法化类型衰落和新的语法化类型兴起的交替。

新旧类型的语法化演变，可能引起语言结构类型的变化。在语言发展的长河中，华夏汉语经历了语法化类型的变迁或语言结构类型的转变。

（一）孤立结构是形态脱落的结果

历史上的华夏汉语是粘着，可能还存在部分屈折的结构类型，在持久的语言混合中形态纷纷脱落，新的语法化过程导致华夏汉语向孤立结构方向持续发展。

**1. 依据历史音系结构，推定华夏汉语具有形态变化**

根据历史语音学研究，先秦汉语是非声调语言，而可能存在复辅音和形态性的辅音尾。虽然汉语的复辅音在东汉已经开始脱落，但根据梵汉对音，唐代可能还存在部分复辅音。若根据音系学的研究，复辅音消失的最晚时间可能延续到北宋[1]。至于汉语的声调，可能在汉代才开始产生，魏晋后才趋于定型。声调的形成受控于语音结构的调节或补偿，深层的缘起可能是语言结构类型的转变。

**2. 依据音素交替子遗，推定华夏汉语具有形态变化**

古代汉字的谐声反映了周秦汉语的屈折变化。古代汉语存在可使字词的语法功能发生变化的-s 尾，所谓汉代的四声辨义即包括去声-s 尾的语法功能。一些汉语方言中，现在依然存在通过声调屈折以表示人称代词的数范畴（陕西商县话）或格范畴（湖南衡山话）。

**3 依据语言亲属关系，推定华夏汉语具有形态变化**

汉语与藏缅语貌异而源近，汉语与壮侗语貌似而源远[2]。

---

① 王洪君《非线性音系学》，北京大学出版社 1998 年。

② 戴庆厦《从藏缅语看壮侗语与汉语的关系》，《汉语与少数民族语言关系研究》，《中央民族学院学报》1990 年 11 月增刊。

现代藏缅语是粘着类型，古藏语(7～9 世纪)形态十分丰富。汉语与藏缅语的“貌异”在于现在的语言结构类型不同，而“源近”的原因在于约公元前 1000 年周族带到中原的是原始藏缅语。现代壮侗语是孤立类型，虽然仍然残留着一些声调交替手段。汉语与壮侗语的“貌似”在于语言结构类型的相同(4 世纪开始，汉语对壮侗语持续渗透)，而“源远”的原因在于前 1500 年～前 1000 年的商族语言可能是原始南岛语。

**4. 依据语言演变规律，推定华夏汉语具有形态变化**

语言学上有一条规律，越是混合的语言其形态脱落得越多，孤立型的汉语可以推定为从形态型的华夏语蜕变而来。英语的演变也说明了这一点，由于历史上多次语言的混合，导致英语由综合型结构向分析型结构发展。换而言之，词法层面上的形态失落和范畴消亡，必须依赖于句法层面上的语序和功能词的补偿。从高本汉开始，赞同或论证古代汉语是形态语言的欧美学者很多。正如契科巴瓦说：“文献确定古汉语并不像现代汉语似的接近于词根语的类型，汉语的词曾经有过附加成分。汉语之接近词根语是长久发展的结果。汉语当日也不是词根语，而是最近几个世纪以来却有力地朝着这一方向移近。”[①]

**5. 依据周边毗邻语言，推定华夏汉语是粘着结构类型**

除了汉语，以及 4～5 世纪前后受汉语巨大影响的苗瑶、壮侗和越南语等，孤立结构类型语区周边的毗邻语言——北方阿尔泰语、西部藏缅语、南部的南亚—南岛语和东部朝鲜语和日本语——都是粘着结构。因此李葆嘉提出原始华夏语是粘着类型的假说。

---

① ［苏俄］A.B.Чикокава《语言学概论》(周嘉桂、高名凯译)，179 页，高等教育出版社 1954 年。

(二)虚字功能是形态范畴的代偿

在形态语言演变成非形态语言的过程中,一方面形态的脱落导致旧有语法范畴的消失,另一方面又以新的方式进行语法化。汉语中的非形态语法化的主要手段,一是词汇手段,通过实词语义来明确句法结构的含义;二是汉藏语中的古老重叠手段的使用;三是与粘着语词缀具有相通性质的虚字(包括语气词、指代词、量词、虚化性副词、虚化性助动词、助词、介词、方位词等)的产生。

迄今为止,汉藏苗瑶语中都没有发现"性"范畴或其残留。汉语的"名量"范畴描绘的是事物的形象特征,准确的说法是"形貌"范畴或特征类别。"形貌"范畴和"性"范畴都是名词的属性类别范畴。汉语量词或类别词的起源包括两个问题,一是汉语[(数+量)+名]结构的来源,一是秦汉以后某一量词的来源。有些学者认为汉语[(数+量)+名]来源于名词回响("羌十羌"——[名+数+名]结构),并且有人举出藏语中也有极类似形式,但是古藏语既没有[(数+量)+名]结构,也没有纯粹的量词成员。然而,库页岛吉俩克语表现在数词和名词的一致上的类别范畴,却发人深省,因为这种[(数+类别)+名词]结构和汉语的[(数+量)+名]结构基本相同。也许,汉语的[(数+量)+名]结构不是起源于[名+数+名]结构,而是起源于是像先秦时代中的"一介臣"这样的结构。在周边语言中,量词最多的是壮语,藏缅语没有类别词,朝鲜语、日本语、艾伊努语、马来语、崩龙语等等却有类别词。如朝鲜语:

| | | | |
|---|---|---|---|
| 人物类 | nat(面) | 动物类 | mari(头) |
| 有柄类 | caro(柄) | 书籍 | kil(册) |
| 米 | al(粒) | 豆腐 | mo(块) |

根据语音对应规则，显而易见，这些类别词不是汉语量词的借用。朝鲜语的底层是南岛语，同时有些日本学者也认为日本语源于南岛语。艾伊努语和吉俩克语，我以为是远古南岛语网向北延伸而断裂的语珠，因此"形貌"范畴或类别结构也许是原始南岛语的特征之一。

假如个体量词作为"形貌"范畴的同时隐含了名词"单数"的规约，汉语中的量词重叠和集体量词似乎隐含着名词"多数"范畴的补偿。至于"双"、"对"、"副"，可以视为远古"双数"范畴的孑遗。随着十进制的成熟，名词"数"范畴形同虚设，最为常见的是以限定性数词作为名词"数"范畴的补偿。在人称代词中，上古汉语存在"侪、等、辈、曹、属"等后缀。唐代出现了词尾"伟"，作为人称代词"数"的补偿，出现在名词前的介词或前置词，出现在名词后的方位词或后置词，都可以看作名词"格"范畴的代偿或现代汉语"格"标系统。古代汉语有表时间、方位的"于"，表工具、目的、方式的"以"字和表领属的"之"等。现代汉语的介词非常丰富，如表时间、表处所、表方向、表方式、表原因、表对象、表比较、表受事等。时态副词、时体助词的出现作为"时体"范畴的补偿。使动式的"使"类字、被动式的"被"类字作为"态"范畴的代偿，"叫"和"让"可以通用。表愿望的能愿词、表祈使的语气词和表假设的关联词语等，作为"式"范畴的代偿。形容词重叠式的使用和程度副词的出现，则成为"级"范畴的代偿。

汉语语法化的单位或手段大致有三个来源：1.华夏语的古老粘着成分，如"于"、"以"等。2.汉语实字的虚化成分，如能愿助词、时态副词等。3.周边非汉语言的借用成分，如时体后缀

"了"、"着"和比况后缀"似"[1]，则是阿尔泰语粘着结构影响的结果。复数后缀"们"（唐代：伟、弭、卫；宋代：满、瞒、们、门；元明：每、们）据研究来自唐代西北方言，唐代西北方言可能是胡人换用汉语后形成的汉儿言语，写法的多样反映了这些不同形体的字只是一个记音符号。元代文献中"们"可作为动物的复数后缀，现在甘肃方言可作为无生命事物的复数后缀，都透露出与阿尔泰语复数范畴的关系。时体与比况后缀，虽然为借用成分但仍可纳入汉语词类而称之为助词。与之相比，"们"的无家可归似乎更显著地表明了其外来身份。

（三）分布功能是语义配比的稳定

秦汉以来，汉语实字的语义配比功能的明确化也在发展中。一方面，秦汉汉语中的使动、意动等"实字活用"，也许魏晋以前还有相应的、但汉字字形反映不出来的形态变化。另一方面，古代汉语中的一些常用搭配或语义结构格式传承下来，形成了实字的语义配比功能。所谓实字的语法功能、分布特征或"广义形态"，其本质就是实字在历史上形成的稳定性配比关系。古汉语中的某一实字，当时具有几种稳定性配比关系的，如只传承下来一种就形成后世的实字定类；如传承下来不止一种就形成后世的实字兼类。某个实字未能流传下来的、非稳定性配比关系，则被后人称为古汉语的字类活用；某个实字未能逐步扩展开来的临时配比关系，则被后人称为古汉语的修辞语用。汉语实字的结构配比是一个历史的动态过程，没有形态变化的管束致使汉语语义配比充满柔性。所谓"词无定类"、"依句辨品"，其本质就是实字的语法功能在习惯性和灵活性、稳定性和可变性的交织

[1] 详见宋金兰《汉语助词"了、着"与阿尔泰诸语言的关系》，《民族语文》1991年第6期；江蓝生《助词"似的"的语法意义及其来源》，《中国语文》1992年第6期。

中此消彼长。尽管为了明确句法结构中的实字功能出现了“的”、“地”、“得”这些形式标志，但是与西方语言的形态变化不可同日而语。因此根据汉语的语感，汉语词类只有根据语义类别进行划分。

## 七、语言的音义二分系统以及传统语法学的位置

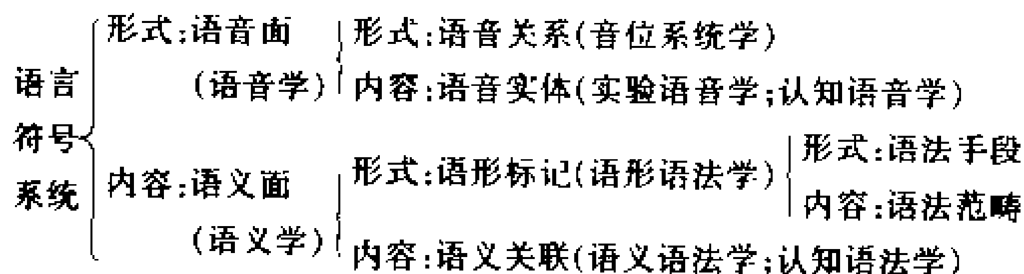
人类建立在经验框架上的认知手段，在语言系统中表现为对象的符号化、类别的范畴化和配比的关联化。语法范畴原型是先民通过语音形式所标记的原始认知语义范畴，形态标记性和类别分布性都是语义结构的句法投影，所谓“概括性的语法意义”是原始认知语义范畴逐步虚化，以致于后人不了解其来源而导致的误解。任何语言都经历过认知语义范畴化阶段，但是语言演变的历史表明，形态手段可能弱化、混化、虚化或完全脱落，以致所表示的范畴意义捉摸不定直至不复存在，从而出现语言结构从形态类向非形态型的蜕变。这一蜕变表明，形态范畴与形态手段并非语言系统的必要成分，语义结构之法是语言的根本之法。

综合结构类型以语音交替、词尾或词缀附加为形态型语形，分析结构类型以语序、虚词和语调为非形态型语形。在形态标记脱落和语法范畴消逝的过程中，语言中会萌发新的代偿成分即出现继起语法化，新的代偿成分和功能范畴形成新的语形系统。任何人类语言都存在两套编码系统，基础性语义编码系统和附属性语形编码系统，但不同结构类型的编码程序侧重可能不同。综合语的语义编码似乎是不言自明，注重的是形态变化编码程序，因此可命名为语形型语言。分析语的语形成分编码具有伸缩性，语义编码程序决定了句子的是否合格，因此可命名为语义型语言。附着在词干上的语法形态具有普遍性或刚性，

语形型语言的研究传统受控于语法范畴,依据形态词类和句子成分的对应性,西洋语法学形成形态词法学和逻辑句法学的二级体系。语义型语言的研究传统注重实字义类和虚字功能,自然形成以实字义类(训诂学)和虚字功能(虚字学)为两翼的语法学体系。

既然语言系统通常被切割为语音、词汇和语法三部分,那么个别语言学研究也就被相应地划分为三个分支学科。然而,语言是语音和语义相结合的符号系统,根本没有"语法"的位置,或者说只有语音的结构法则和语义的结构法则。在具有形式和内容关系的语音和语义的二分系统中,传统语法的位置何在?语义面同样可以区分出内容和形式,其内容是语义结构,其形式即语义结构的投影——语形结构。依据这一观点,以语形为主要研究对象的"语法学"是语形语法学,以语义为主要研究对象的"语法学"是语义语法学。

**语言的音义二分系统以及相关学科**



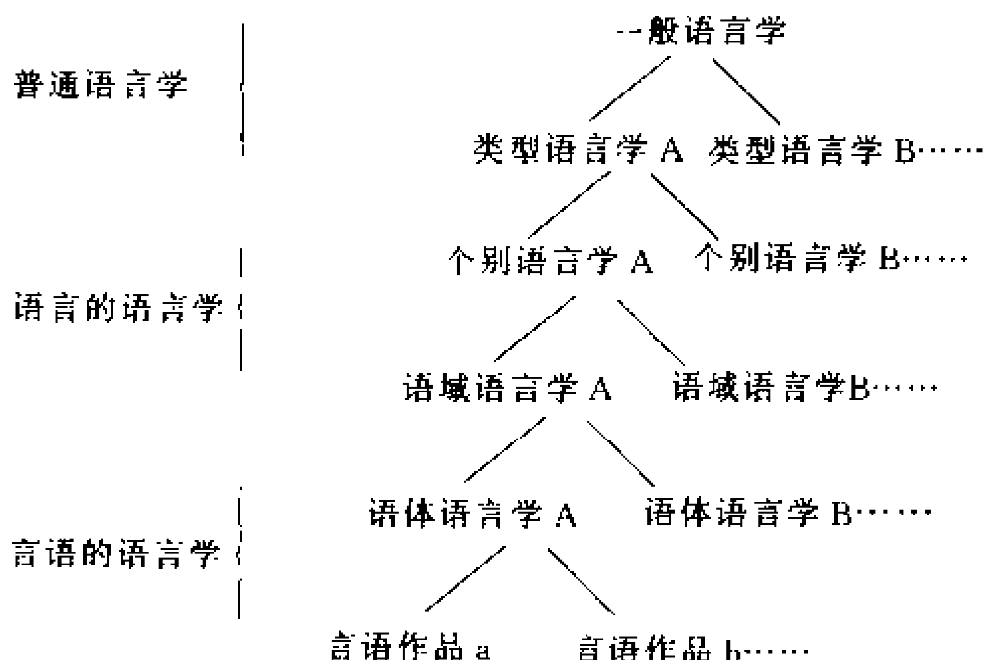

传统语法研究以形态范畴为对象,划分为语法形式和语法意义,研究重点在语法形式,语法意义被抽象为脱离语义的形式结构意义,从而形成了语音、语义和语法的三分错位,凸现了作为虚象的形式结构,而掩盖了实相的语义结构。传统语法研究的是表层形式系统,在语言工程学中,语形语法不仅对句法结构的解

释力微小，而且依据语形规则生成出来的句法结构其正确率极低。既然依据形态语法范畴所归纳出的语形结构系统，不是自然语言有效生成型的自足语法系统，那么应当依据现代语言事实重新归纳认知语义范畴，建构生成型自足语义语法系统。揭示形态语法范畴的原始认知范畴的原型或语义类别的本质属性，正是为了建立语义语法学，进一步推动语义结构网络的研究，以促进自然语言的计算机处理。

从洪堡特创立普通语言学以来，欧洲学者的语言视野逐步扩大，具有普遍意义的语言理论日益受到重视。但是，不得不看到，一方面，西方学者所提出的理论主要建立在欧美语料或诸多形态语言的语料之上，甚至坚持运用以英语语料为对象就可以研究出人类的普遍性语言能力；另一方面，而对存有隔膜的西方语言学理论，中国的一些学人曾经提出建立具有“中国特色”的语言学理论。尽管从普遍唯理语法学派开始，西方学者就追求“普遍”，但是迄今为止，“普通语言学”其实并不那么普通，距离世界普通语言学的目标依然存在着很大距离。

面对“中国特色”，有人质疑——莫非每一种语言都要建立一种语言学理论？那么世界上几千种语言不是就要建立几千种语言理论。毫无疑问，如果可能，每种语言都应当建立自己的个别语言学，每一个别语言学都是描写和理论的结合。只要是不同的语言，自然具有自己的个性，就可以抽象出相应的理论。但是，普通语言学的研究，并非要等到所有的个别语言学都建立以后才能着手。我们可以首先考虑世界语言学研究的层级体系，然后在这一框架内展开的研究就可能具有预期性。

世界语言学研究的层级体系




由此观之:1.言语作品是社会成员对语言系统的具体使用,语言系统是语言学家对言语作品中的抽象建构,言语作品是语言学研究的学科对象或材料来源。2.由于社会成员的差异性和社会情景的不同性,根据相应的表达要求表现出不同的语体风格,依据这些材料研究的结果就是语体语言学。3.对同一语域中的不同语体语言学进行再抽象,其系统就是语域语言学。4.在不同语域语言学的基础上再抽象,其系统就是个别语言学。5.将相同类型的个别语言学进行继续抽象的结果就是类型语言学。6.以不同的语言类型学为基础进一步宏观抽象,其结果才是一般语言学。

言语的语言学是微观研究,包括个人言语作品研究和语体语言学。语言的语言学是中观研究,包括语域语言学和个别语言学。个别语言的阶段性研究是共时语言学,个别语言的不同

阶段性研究是历时语言学。对个别语言不同地区性语言研究就是地区方言学，个别语言的不同社会性语言研究就是社会方言学。普通语言学是宏观研究，包括类型语言学和一般语言学。依据语言演化方式的异同，可以建构语言关系类型学（语网、语珠、语区、语系）；依据语言结构方式的异同，可以建构语言结构类型学（复综、屈折、粘着、词根）。一般语言学所研究的是人类语言的普遍共性，建立一般语言学的前提是类型语言学。人类语言系统的本质共性可以概括为——建立在像似性认知中的、以语音为手段的语义性，因此语义语法学才具有普通语言学的真正价值。

## 第二节　面向信息处理的语义语法学理论

### 一、语言学流派和语法学目标的嬗变

（一）语言的结构类型与自发的研究模式

雅可布逊说："语言代码的整个模式，在很大程度上决定着他们对语音的听觉感知"，而听觉感知决定了语主对其母语语音的分析模式，这一直觉分析模式就包含在文字的类型或音符的性质之中。语言的语义结构和形态变化都要通过语音来表示，语音要素的感知与语言结构的分析具有一致性，由此导致了语法的自发研究模式。因此依据文字类型，可以看出语言类型和语法的自发研究模式之间的制约关系。

粘着结构的特征和记录词缀的需要，促使苏美尔文字中产生了人类历史上最早的音节符号，因此粘着语的研究从词根和词缀的分析开始。虽然粘着语的自发研究无疑最早，但迄今为止苏美尔文献中有没有语法书之类的泥版文书未详。而且即使

存在,其方法大概也不会被语言结构不同的巴比伦文化承传下来。对粘着型的突厥语的研究,最早是马合木德·喀什噶里的《突厥语词典》(11 世纪 70 年代)和《突厥语句法精华》(已失传)。由于他精通突厥语和阿拉伯语,因此对突厥语的语法研究必然受到阿拉伯语法学的影响。

屈折结构的闪含语言,其辅音词干固定而元音交替的特征,促使古埃及文字中产生了人类历史上最早的音素符号。经过闪米特人的借用形成辅音字母,希腊人又在腓尼基字母的基础上增加了元音字母,因此屈折语的研究从表示形态变化的音素分析开始。屈折语所具有的丰富形态变化成为词类划分的天然标志,决定了古希腊语法研究的特色是形态分析法。同样是屈折结构的阿拉伯语的语法研究,在 8 世纪的西巴维希(Sibawaihi)的《书》(*Al-Kitabu*)中趋于成熟。西巴维希吸收了语言结构相近的印度和希腊语法体系,对阿拉伯语特有的三辅音词干的描写成为最重要的贡献。

与西方形态语言不同,记录华夏汉语的文字在象形的基础上产生了假借,假借的广泛使用势必导致形声的出现,向孤立结构演变的华夏汉语的研究从假借的语义探求和功能辨析开始,因此中国古代最早出现的是聚合字义的义类辞书《尔雅》。面对"邦言"的复杂性,扬雄的《方言》是一部对比词汇学著作。依据汉字的系统性,许慎的《说文》分析字形以求本义。刘熙则探索"名之于实,各有义类",其《释名》成为一部语源学名著。唐宋以来,在秦汉典籍的阅读和研究之中,导致了古音学的昌明和虚字学的成熟。元代学者卢以纬撰写成中国第一部虚字学专著《语助》,后世学者注重虚字的句法功能研究。明清学者提出"欲通古义,先通古音"、"文字训诂,本乎声音",一再强调音义的相互关系。

20 世纪初,在调查印地安语的过程中,鲍尔斯等人发现了

与欧洲语言结构不同的复综语，提出了必须依据语言结构特征来制订研究方法的观点。而对没有“词”而只有“句”的印第安语，只能采取逐层切分法。人类学家在学习印第安语时总是先听语音再学语句，由此他们觉得语言分析从语音到语句的次序天经地义，理论的绝对化表现为制订出针对一切语言研究的“发现程序”。虽然其贡献在于以此揭示了语言结构的层次性，但是存在着把研究过程混同于学习过程的缺失。

如果语言的自发研究最早体现为自源文字的创制，那么苏美尔文字的发明就是世界语言学的渊源，而古埃及文字中的离素原则成为屈折语语法研究的滥觞，腓尼基字母的发明为希腊语法研究提供了必要工具。粘着结构导致词根和词缀分析法的自发产生，形成词根和词缀研究的语法体系；屈折结构导致形态分析法的自发产生，形成词法和句法的语法学二级体系；孤立结构导致实字语义和虚字功能的自发研究，形成训诂学和虚字学的两翼传统；而针对印地安语复综结构的特点，语言学家才提出了层次分析法。由此观之，中国不可能拥有西方的形态语法学传统，如果从人类普遍语法学的立场出发，不拘泥于“印欧语眼光”，可一言以蔽之——建立在汉语直觉感知基础上的训诂学和虚字学就是中国的传统语法学或语义语法学传统。

（二）语言观的变革与语言学流派的形成

西方智者在对客观世界和精神世界之间的中介——语言的思考中，孕育了希腊的哲学语言学；先秦哲人在天道人伦的思考及古代经典的传承释读中，诞生了中国的经学之小学。就西欧语言学而言，虽然逐步形成了相对独立的形态语法学方法，然而相关学科理论方法的借鉴往往导致语言理论的嬗变。一方面，语言的世界是一个万花筒，不同的审视角度则形成不同的语言观，另一方面，语言的研究者学有专攻，不同的知识结构则形成

不同的语言观。无论是审视角度的不同,还是知识结构的差异,不同的语言观则必然孕育不同的语言学理论或语言学流派。

**1. 语言基本工具论**

在醉心于探索意识与世界的希腊哲人眼中,语言是思维工具,语言分析、哲学思辩、逻辑规则混成一体。形态变化的发现,促使形态归纳实证法的产生以及词类系统的构建,最终导致了前1世纪的希腊形态词法学的成熟[①]。虽然在形态丰富的语言中,词法学研究已经解决了造句规则,但是由于语言形态的脱落或逻辑主义的倾向,肇始于2世纪的句法学[②]在13世纪的摩狄斯泰(Modistae)学派的思辩语法(speculative grammar)中才日臻成熟。一方面,拉丁传统语法一直被作为阅读和交际工具,语法始终处于中世纪的"七艺"之首;另一方面,为寻找语法的哲学基础而促使逻辑学家和语法学家的合作,波尔·罗瓦雅尔(Port Royal)学派中出现了普遍唯理语法。显而易见,欧洲语法学与哲学、逻辑学始终结下了不解之缘。

与西方的思维工具论或泛逻辑主义主导有所不同,在印度和中国学者眼中,语言是阅读典籍的工具,语言研究的目标就是为注释经典文献服务。公元前4世纪的印度巴尼尼(Panini)《梵语语法》(*Astadhyayi*)的内涵是语言结构法则,而不是20世纪的"窄化语法",包括词法学、语音学等。中国传统小学研究的对象是五经等文献,以经学道统为基础而附庸于经学的小学语言观和系联法,许慎阐述得尤其清楚。《说文·叙》曰:"盖文字者,经艺之本,王政之始。前人所以垂后,后人所以识古。故曰本立而道生",此为小学语言观。《说文·后叙》曰:"方以类聚,物

① 特拉克斯(Dionysius Thrax)的《语法学》(*Techne Grammatike*)

② 狄斯科鲁斯(Apollonius Dyscolus)的《论句法》(*Peri Syntakseos*)。

以群分，同条牵属，共理相贯，杂而不越，据形系联，引而申之，以究万源"，此为小学方法论。具有朴素系统论的系联法贯穿于中国语言研究的各个方面，《尔雅》《方言》是字义系联，《说文》是字形系联，《释名》是音义系联，韵书是反切系联，虚字研究则是功能系联。传统小学方法当界定为"系统·系联"法，其哲学观是与汉语音节模式同构的整体论和主体心象论；西方语言学的传统方法是"分析·归纳"法，其哲学观是与字母音素模式同构的原子论和客体镜象论。

2、语言生物机体论

19世纪的欧洲语言学家，在发现印度梵语和欧洲语言相似性的基础上，转向历史比较语言学。历史比较语言学的研究方法是语言比较，研究目标是祖语构拟，研究成果是语系确立。在生物进化论的影响下，施莱歇尔把语言看作生物体，在语言生物主义的基础上建立了谱系树模式。一些欧洲语言学家进一步提出了语言类型进化模式和文字类型进化模式，在生物机体语言观的导向下，却充分曝露出盲目的"欧洲中心主义"和狭隘的"印欧语眼光"。

3、语言结构系统论

作为历史语言学的叛逆，在索绪尔看来，语言学的主要任务就是研究静态结构系统内的组合关系与聚合关系。在语言是一个同质性符号系统和语言价值不在于实体而在于形式的格式塔(Gestalt)语言观的主导下，采用的是以社会心理学为哲学基础的实证主义归纳法。反对纯正结构主义的布拉格音位学派，建立了结构和功能、共时和历时相结合的方法，而哥本哈根语符学派则采取高度抽象研究法，发展了符号学理论。以行为主义心理学为基础的美国描写主义绝对排除语言的意义，从而把结构主义的形式化研究推向极端。法国语言学家在赞同语言是结构

系统的同时,不主张"结构主义"和"单纯形式",由于重视语言的功能而形成法国结构—功能学派。最有影响的成果就是特斯尼耶尔(L. Tesniere)创立的从属关系语法,由此孕育而在德国语言学家那里成熟的配价语法学理论与方法。

**4. 语言文化形态论**

比同时代任何一位欧洲语言学家的视野都要宽阔得多的洪堡特,冲破狭隘的印欧语眼光,在创立普通语言学同时确立了人文主义语言学思想,提出了"语言是精神创造"和"语言世界观"的论断。20世纪初的人类语言学家,无论是鲍尔斯(F. Boas),还是马林诺夫斯基,都认为语言是文化形态或精神创造。离开语主的文化或精神世界,对于南太平洋群岛语言或美洲印地安语言的现象则无法理解。萨丕尔和沃尔夫(Whorf)进一步提出"语言相对论",强调了语言系统和意识形态的互塑作用。人类语言学研究采取的是人类学方法和文化阐释法。

**5. 语言行为方式论**

马林诺夫斯基在《原始语言中的意义问题》(1923)中认为,应当通过一个民族的文化生活和风俗习惯来观察语言的功能,与其把语言看作"思想符号"不如认为它是"行为方式"。弗斯(J. R. Firth)接受了马林诺夫斯基人类文化学的"情境说",认为语言是一种社会过程或人类行为,形成了情境语言学或功能语言学学派,采取的是情境解释法。50年代,斯特劳森(P. F. Strawson)在《论所指》中对语言行为的研究开始细化,第一次区分了"语句"、"语句使用"和"语句讲说"。英国哲学家奥斯汀(J.L. Austin)在《行为动性和讲述性》、《如何用词做事》中认为,语言可以分为讲述性语言和行为性语言,又把语言行为分为辞面行为、辞内行为和辞后行为。

6. 语言社会现象论

社会语言学则认为语言是社会的语言，社会性是语言的根本属性之一。语言的变化不仅与语言内部规律有关，而且与社会密切相关。语言是一种社会现象或社会关系，语言社团的差别导致了社会方言的形成。在研究中采取了社会学研究中的调查对比法和统计分析法，由此形成了以拉波夫为代表的微观社会语言学学派（城市方言学派）和以费什曼(J. Fishman)为代表的宏观社会语言学学派。

7. 语言先天能力论

欧洲17世纪的普遍唯理语法认为语言规则是人类普遍理性的表现，试图寻求"语言艺术的自然基础"和"适用于一切语言的一般原则"，采取的是以唯理主义哲学和逻辑为基础的演绎法。20世纪50年代出现的生成语言学或形式语法学，进一步认定语言是人类的先天能力，试图发现与生俱来的语言获得装置(LAD)。乔姆斯基强调语言学属于自然科学，采取的是以唯理主义哲学和数理逻辑为基础的演绎法。

8. 语言认知能力论

皮亚杰(Piaget)曾经提出，语言能力是认知能力的一部分。认知和语言的密切关系以及相关学科合流的趋势，必然导致认知语言学的产生。与社会语言学从社会意蕴、功能语言学从交际功能有别，20世纪80年代出现的认知语言学派从普遍认知能力出发，寻找隐藏在语言现象之中的认知理据，以兰噶克尔(R. Langacker)、菲尔默(Fillmore)、拉考夫(Lakoff)、克饶夫(W. Croft)、海曼(Haiman)等前生成语言学家为主而形成"后乔姆斯基学派"。

认知语言学认为语言是一种基于人类一般认知能力、负载交际功能的符号系统，试图将语言内部形式规律诉诸非形式化

的直觉性外在解释。其基本理论可以大致归纳为:(1)语言的性质,反对"语言是逻辑的、可计算的"的形式语言学观点,认为自然语言是人类认知活动的产物和工具,其结构和功能应视为人类一般认知活动的反映结果。人类的语言能力并不是一个独立于人的其他认知能力和知识范围而完全自主自足的天赋部分,而是与其他认知能力密切相关的普遍认知能力的重要组成部分。(2)语言系统的性质,与形式语言学认为语言系统是一种能生成仅在语法上正确的句子的自主机制,无需参照其他认知因素不同,认知语言学认为语言系统是整个心理组织的有机组成部分,在具体运用中能触发其他认知系统。(3)句法结构的性质,句法结构并不像生成语法学家声称的那样是自主的,句法仅仅是向说话人提供一份清单,说话人必须依靠自己的认知能力,结合语义、语用等句法之外的因素造出合适的表达式,句法结构与语义结构对应。(4)语义的性质,认知语言学尤其注重语义,语义并不限于语言系统内部,而是植根于说话人的知识系统和信仰系统。语义具有主观性,纯语义知识和百科知识并不能截然分开。意义是一个植根于知识网络和信仰系统中的认知结构,理解一个语言形式的意义必须激发相关认知领域中的其他认知结构。(5)语言的解释,与人类具有的共同认知能力对应的语言存在普遍共性,但语言共性往往体现为某种趋势,不一定是绝对规则。对语言共性更有意义的解释往往须在形式之外结合语义、交际原则以及认知能力和认知策略来综合考察。

认知语言学的基本理论建立在人类经验基础之上,客体实在性决定了认知理据性。经验基础是人类在对世界的长期认知活动中所形成的特定的思维图式和观念心态,既包括人类普遍的认知心理,又包括不同社团的集体共有的文化背景和个体特有的知识背景。语言的发展以最初的经验基础为前提,人类的

基本认知能力包括时空化、范畴化、网络化、凸显化、隐喻化等认知能力。语言符号与外在世界和人类心智的密切关系，突出表现在语言系统与现实世界、认知经验的像似性上，这种“语言相”与“世界相”或“经验相”的关系简称为语言的像似性。语法结构的本质是象征性的，是人们现实感知的反映。语言不是一个自足的认知系统，句法也不是一个自足的形式系统。“意象”、“认知域”、“突显”、“象征单位”、“扫描”等是认知语言学中的重要概念。语言现象与现实事态的像似性广泛存在于各种词法、句法组织中，以致于体现为语言共性规律。认知语法学家对以句法象征性为中心的“自然句法”做了大量而深刻的探讨，并逐渐总结出复杂性、独立性、有序性、重叠性、距离性等五个具有普遍意义的句法像似原则。主张建立在经验基础上的语言结构和人所面对的世界图式具有像似性的认知语言学，采取的是以认知心理学为基础的类比解释法。

（三）时代的变迁和语法学目标的嬗变

在古希腊哲人的主观意识和客观对象的关系探索中，语词的研究源于对客观对象真谛的争论，句类的划分源于修辞学家对美文技巧的分析，句子的分析源于思辩学家对表达逻辑的比对。词类的划分，从柏拉图含混的二分，经过亚里斯多德的三分，再到斯多噶学派的五分，由此形成了语词概念意义和形态变化相互印证的词法学。虽然伴随着逻辑学的发展，逐步形成句子成分的概念，但是直到思辩语法对主语和谓语的明确，西方词法和句法的语法二级体系才最终完善。拉丁语语法教学促使规范语法日益显著并成为语文应用的工具，13 世纪经院哲学家的思辩语法才把欧洲语法研究提高到理论层面。沿着思辩语法的道路，17 世纪的法国波尔·罗瓦雅尔学派的逻辑学家和语法学家共同建立了普遍唯理语法理论。19 世纪的理论语言学成就，

首推历史语言学的比较研究法和谱系树模式,其间洪堡特等尝试建立的普通语言学理论和语言结构类型学,从印欧语的狭隘视野转向世界诸语言。20世纪的人类语言学、结构语言学、功能语言学、社会语言学、生成语言学、文化语言学和认知语言学,无一不受惠于洪堡特的普通语言学。

与西方语言学的嬗变不同,源远流长的中国小学,前后形成训诂学、对比词汇学、文字学、语源学、音韵学和虚字学研究的六个分支。一方面,具有以雅学、说文学、注释学为主流的语义阐释学;另一方面,具有重在虚字功能的语形虚字学。传统西语语法学突出词法的形态性、句法的逻辑性和表达的规范性,而传统汉语语法学突出语义的阐释性、虚字的功能性和经学的实用性。随着19世纪历史比较语言学的建立,印度和欧洲的语言学传统实现了合流。19世纪末,《马氏文通》的出版标志着东西方语言学传统的合流拉开帷幕。经过一个世纪对西方语法理论和方法的模仿,由于语言类型的巨大差异,以西方语形语法学为基点的合流步履维艰,实际上已经流产。反思的结果是必须重新寻找东西方语言学传统合流的新基点——从语形语法学转向语义语法学。

引人注意的是,自然语言的计算机处理对西方语形语法学提出了根本性的挑战,也为东西方语言学传统的合流带来了新的契机。西方传统语形语法学面对的是人与人之间所进行的语言交际活动,作为语言的创造和使用者都天然具有建立在认知基础上的语义网络,为语言活动提供了认知的、社会的、文化的背景。在自然语言的计算机处理或人机对话中,人所面对的是不具有建立在认知基础上的语义网络的计算机,排斥语义结构研究的语形语法学顿时陷入困境。可以认为,几千年来的西方语法学或语形语法学研究之所以能够延续下来,是因为语言的

创造和使用者的语义网络和认知能力"先天"存在。由于这一语义网络系统的"先天"存在，致使以往的语言学家误以为语形语法就是语法的全部，而心安理得地认为语义研究不是语法学家的任务。

尽管《句法结构》的出版被喻为"乔姆斯基革命"，但是生成语义学才是语言研究本体的彻底转向，在形式语法学步履维艰的情况下，西方相继出现蒙德鸠语法、词汇功能语法、优选语义学和语义网络理论等。在中国，无论是配价语法和格语法的引进，还是立足于汉语的句法语义学、意合语法学的倡导，都促使从语义角度阐释语法结构的研究成为世纪末的热点，然而，迄今为止，汉语的计算机处理应用系统自觉而全面地运用语义研究成果的几乎没有。虽然其表象似乎是语义研究现有成果不多或未成系统，但其根源却在于以往的语义分析研究项目，没有清醒认识到——必须从人类的（不限于印欧种群的）、普通的（非欧洲唯理主义的）语言学（非形态语言的）的高度，来全面解决理论的彻底创新和方法的具体而微。第五代计算机研制的搁浅表明，语形语法学无法解决自然语言的计算机处理。无论是自然语言的计算机处理，还是语言研究的进一步深入，汉语语言学都需要既能反映汉语典型个性又植根人类语言本质共性的原创理论。

## 二、人类语言的本质共性是语义性

### （一）形态语法范畴的本质是语义类别

人类建立在经验框架上的认知手段，在语言系统中表现为对象的符号化（单位）、范畴化（类别）和关联化（层级性配比）。形态语法范畴的原型是先民通过语音形式所表现的原始认知范畴或语义类别，形态标记性、结构层次性和单位分布性都是语义类别关联化的句法结构投影。先民通过经验框架观察世界图式

并将之语言符号化,这一过程表现为世界图式——经验框架——语义范畴——语形标志。从世界图式到语义世界,一方面由于有相同的感觉器官和大脑机能,不同的语言文化种群得到的世界图式大体一致,另一方面不同的语言文化种群借助各自的经验框架获得的语义范畴,逐步语法化为不同的语形标志。无论是原始语言中的认知语义范畴,还是后世语言中的语形标志,都是先民范畴化认知的结晶,语形范畴只不过是语义范畴逐步蜕变的表层标志。以原始认知语义范畴为基础的这些表层语形,一方面不能涵盖以后形成的认知语义范畴,另一方面又不可避免地出现混化、弱化,以致于表层语形的完全丧失而最终导致语言结构类型的嬗变。形态语言可以转化为非形态语言学的事实证明,形态语法范畴并非语言符号系统的必要部分,原始形态语法范畴可以被后起的语形范畴所替代,也可以由词汇语义来表达。

(二)不同语言可沟通的基础是语义性

不同的语言能够相互理解和翻译,不在于语形结构的差异性,而在于语义结构的沟通性。不同的语言的理解和翻译过程,就是排除语形差异性和依靠语义沟通性的过程。把A语言翻译为B语言的基本过程,就是脱下A语言的语形服装,将A语言的语义转达为B语言的语义,然后再穿上B语言的语形时装。语言翻译过程中,最棘手的问题不在于语形的是否取舍,而在于语义的如何转达,因为语义的转达牵涉到双方的语义场、语义结构和语言的文化功能和联想心理。即使最严密的语形规则系统,也并非语言符号生成转换的自足系统,反之,一种形态语言中的语形语法意义,完全可以转化为另一语言的词汇意义。归根结底,语言的结构法则是语义结构之法。电脑需要的正是具有人脑语言共性的、可以表现为不同自然语言互译基础的语

义网络的程序化系统。

（三）古希腊语法研究肇始于语义探讨

语言作为人类最重要的认知方式、文化形态、交际工具和符号系统，是人类文化的综合体，对语言的不同方面的侧重则产生不同的研究走向。并非像很多人所认为的那样，西洋语法学一开始就自觉从形态入手。一方面，希腊语法研究肇始于客观世界和精神世界中介的语言意义的哲学思考，形态发现后，才从语形和语义的印证转移到词类形态划分和句法逻辑阐释，以致于"买椟还珠"。另一方面，主张意义和形式相结合、与形式研究彼消此长的意义研究时断时续，就是在彻底摈弃语义研究的结构主义盛行的时代，西方仍然存在"从形式到意义，从意义到形式"的叶斯泊森（新传统语法学和强调意义的英法功能语言学）。60年代以后，面对形式结构规则生成的句子破绽百出、不忍卒读，乔姆斯基不得不把语义纳人语法，先后提出深层结构和表层结构的概念，尽管进一步发展为解释语义学，但是语义结构研究仍然是"宗教的婢女"。直到麦考莱（J.P.Mecawley）等人提出语义规则是句法规则的基础的生成语义学，才实现了语言法则研究的根本性转向。

（四）人类语言系统的本质共性观

人类语言都有语义编码与语形编码两套程序。尽管可以依据语码的结构特点和编码的程序侧重，将人类语言分为语形型语言和语义型语言，但是无论哪种类型的语言的编码都必须首先启动语义编码，就语形型英语和语义型汉语而言，其区别在于前者务必再进行以形态变化为手段的语形编码，后者务必再进行以完句成分为手段的语形编码。在语形编码过程中，形态丰富的西语的语形落实在词级层面上，所以词法学已经包含了句法学，而没有形态的汉语的完句语形凌驾于句级层面上，所以词

组的结构不等于现实的句子；西语的语形手段具有普遍性和刚性，汉语的语形手段具有调节性和柔性。由此决定了语形型语言的编码程序重在词法或句法表层的形态一致性，而语义型的编码程序重在语词选配或语义深层的结构关联性。因此也可以给这两种结构类型重新定义，西语是词法型语言，语法学的特点是研究形态变化的词法学；汉语是词汇型语言，语法学的特点是研究实词搭配和虚字功能的词汇学。

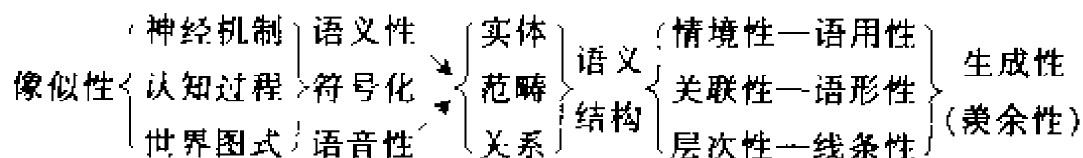
尽管形态型或词法型语言的编码重在表层词法形态，但语义编码仍然是不可忽视的必要程序，因为：1. 词法形态必须落实在语义编码上；2. 错误的语义编码，即使词法形态无误仍然是不正确的句子。“话题—说明”和“主语—谓语”之间的差别，不是句法表达框架层和语法形式结构层的差别，而是两种表达框架或句法结构类型的差别。因此，西语和汉语的不同编码方式层次可以图示为：

| 语言结构层面 | | 语言结构单位、框架与手段 | | |
|---|---|---|---|---|
| 语形层 | 语法形式结构层 | 语形手段 | 完句成分 | 形态变化 |
| | 句法表达框架层 | 表达框架 | 话题—说明 | 主语—谓语 |
| 语义层 | 符号序列语义层 | 语义结构 | 字组（读） | 短语（phrase） |
| | 符号个体语义层 | 语义单位 | 字（名） | 词（word） |
| 认知基础和表达要求 | | | | |

处于句法表述框架层的“主语—谓语”和“话题—说明”，其下层都有相应的符号序列语义层，其上层都具有各自的语法形式结构层。

由此可见，人类语言系统的共性可以概括为：建立在像似性

感受认知模式中、以语音为手段的语义性，以及由此派生的情境性—语用性、关联性—语形性、层次性—线条性、生成性—羡余性等不同属性，可以图示为：




总而言之，语义性是人类语言系统的根本属性，语义性是人类语言系统的其他属性的基础，其他属性都是人类语言的派生或表层属性。根据以上观点，必须重新建构具有普通语言学理论价值的语义语法学。

## 三、语义语法学的语法观

（一）现代汉语是语义型语言

虽然语法意义在形式上部分表现出来的语形型印欧语，可以采用词本位（从语言的静态组成单位出发）语法体系或句本位（从语言的现实结构体出发）语法体系，但是在汉语语法体系的描写中，正确认识汉语的结构类型才是建构适合的语法体系的基础。20世纪的汉语语法学研究，在背离中国传统语法学立场的同时，因为套用西方语形语法学框架而出现了三个显著的基本矛盾：1.西语的词类与句子成分具有对应性而汉语的词类与句子成分具有交错性的矛盾；2.西语的基本语言单位word具有形态性和汉语的基本语言单位"字"具有调节性的矛盾；3.西语的句法表达框架层的"主—谓"框架和汉语的句法表述框架层的"话—说"框架的矛盾。

面对第一个矛盾，朱德熙（1985）提出词组本位说。主要思

路是：划分汉语词类只能根据词的语法功能——层次分析是语法分析不可缺少的手续之一——语法研究的最终目的是弄清语法形式和语法意义之间的对应关系——汉语词类跟句子成分不存在简单的对应关系——汉语句子的构造原则跟词组的构造原则基本一致——词组本位语法体系。根据词组本位说，句子不过是独立的词组，词组和句子之间是实现关系或抽象的结构和具体的话语之间的转化关系。虽然词组本位观抛弃了句本位观中的句子成分概念，避开了句本位观中词类与成分之间的纠缠不清，但是在用词组结构分析代替句子成分分析显得简便的同时，也就掩盖了备用单位和使用单位的本质区别，抹杀了从词组到句子的实现条件。显而易见，构成汉语句子的不是词组，而是作为句干成分的词组和必要的完句充分，所谓完句成分是指依附于实词构成的句干并使之实现为动态句子的过程中所需要的句级功能性成分。以动词谓语句为例，汉语句子构造的层级模式如下：

| | | | |
|---|---|---|---|
| 谓核结构 ……………… | 语义核心 | 句干成分 | 语言符号 |
| ＋状态元 ……………… | 语义扩展 | 句干成分 | 语言符号 |
| ＋词序 ……………… | 分布完句成分 | 语形成分 | 语言符号 |
| ＋虚字 ……………… | 字词完句成分 | 语形成分 | 语言符号 |
| ＋韵律 ……………… | 语音完句成分 | 语形成分 | 语言符号 |
| ＋语境 ……………… | 非语言符号 | | |

“汉语词类跟句子成分不存在简单的对应关系”，这不是汉语自身的特点，而是西语语法框架扭曲汉语语法事实的误会。“汉语句子的构造原则跟词组的构造原则基本一致”，这是汉语自身的特点，但句子跟词组构造原则基本一致的实质是汉语语义结构

规则的一致。

面对第二个矛盾，徐通锵先生提出字本位说，并且在赵元任关于汉语的基本单位是“字”的论述中找到了支持。汉语语义语法学的研究单位是“字”、“字组”（包括凝固字组、稳定字组和即兴字组）和“句子”。在语言单位中，现代汉语的“字组”的构造以具有语音结构和语义功能的单个音节为基本单位。英语的单个音节与语义单位并不对应，只是word中组成音节，语义表达的任务必须依靠多个音节的整体组合去表现。与之不同，现代汉语“字组”的整体语义，由具有语义功能的独立音节联络而成。“字”既是一个音节单位，又是一个语义单位，这才是汉语构造单位的基本特点。因此从基本单位而言，字本位说比词组本位说切合汉语事实。

面对第三个矛盾，赵元任提出话题主语说，20世纪80年代以来，海峡两岸的学者们先后进行了进一步的讨论。英语句子框架使主谓双部，主语名词与谓语动词互相制约并且要求其他成分与之具有一致关系，由此形成内部结构严谨、外部句界分明的表达格式。而汉语句子的结构，既包括加上完句成分的双部表达格式，也包括加上完句成分的名物性、动作性、性质或状态性等的单部表达格式。以没有形态制约性的“字”为编码单位的汉语，由此形成内部结构柔性、外部句界游移并在很大程度上依附于语境的表达格式。西语的“主—谓”框架以限定动词为中心，“主语”是“谓语”的对应性或强制性必有成分，汉语的“话—说”框架既不存在限定动词，也不一定要以动词为中心，“话题”与“说明”具有互补性而不具有强制性，可以互补也可以独现。

只有清醒地认识到汉语是语义型语言，才能彻底摆脱建立在形态语言事实上的词类形式划分和成分逻辑分析，只有彻底摆脱建立在形态语言事实上的词类形式划分和句子逻辑分析，

才能真正回归汉语研究的基点，才能够走出西语语法框架扭曲汉语语言事实的怪圈。汉语语法研究当以“字”为基本单位，实字的语义结构规则系统和虚字的语法功能层级系统——才是汉语语法研究的对象。

（二）语法的投影虚象和真实本体

欧洲“语法”术语的含义并非一直仅指词法或句法。在传统文献语言学研究让位于历史语言学研究的过程中，“语法”这一术语在19世纪也就相应扩展为“语言法则”，因此19世纪末的新派历史比较语言学家才称之为“青年语法学派”。“语法”这一术语的窄化，是20世纪结构主义和描写主义研究对象日益窄化的自然后果。语法即语言结构法则，包括语音结构法则、语义结构法则，以及以语音为手段将语义结构法则投射在表层上所形成的——语形结构法则，还有受制于具体场合的语用结构法则。作为语义码、语形码、语用码载体的语音法则，包括词音系统、音位系统和韵律系统。语义编码是语言编码的内核，语形编码是语义关联的标记化投影，语用编码是语义表达的情境化手段。

生成语法学的创立者乔姆斯基认为，语言反映了人类的明智能力或天赋能力，语言研究的最终目的是揭示人脑和知识本质，但是依据乔姆斯基的逻辑思路：普遍语法（研究对象）——语言能力（哲学基础）——自然主义（学科性质）——数学方法（方法借鉴）——符号描写（形式载体）——形式语法（研究目标），显然缺少人类天赋成分中最关键的语义结构能力这一环节。生成语法学的经典理论阶段完全排斥语义，标准理论阶段不得不引进语义，但仍把句法结构作为基础，语义结构仅作为解释手段，执著地抱住语形的形式语言学的研究旨趣不放。

语义语法学的语言观认为，语言符号是建立在认知范畴化基础上的语义结构系统，语言系统是建立在认知符号实体范畴

化和事理关联化基础上的语义网络。语言的萌芽是源于个别事物或具体事件的单个语言符号的创造和使用，语言符号个体（实体）和语言符号序列（结构）是人类经验基础框架中的世界摹本，语言符号序列不过是符号个体内部结构关系的放大和扩展。人类认知的本质是对世界图式的语义符号化，所谓句法不过是语言符号个体扩大了的、语言符号序列内部语义关系的标志化。因此广义语义，不仅包括实体范畴化的狭义语义，还包括关系范畴化的句法意义和语境范畴化的语用意义。语义编码是一切语言编码的基础，任何语言中的语义编码错误的句子都是不能成立的。语法的真实本体或语法研究的对象应当是语义单位、语义类别及其结构关联。可以这么比喻，以往的形式语法学犹如把观察一个人的影子当作在研究这个人的本身。

"形式语法学"却把语法的投影虚象——语形——作为语法研究的惟一的对象。由于自然语言不是经过高度再抽象的规范化的数学语言，所以任何脱离语义结构的狭义语法的形式化注定流产。乔姆斯基革命仅仅是革了"描写主义"（归纳法的系统描写）的命，结构主义的研究对象"形式"（语形）的命并没有革去，因此乔姆斯基生成语法学和结构主义的共通之处在于以语法的投影虚象语形为研究对象，所不同之处在于结构主义把形式看作语言系统本身，而生成语言学把形式看作人类语言能力。尽管乔姆斯基的形式语法使人们进一步认识到语言形式化的重要性，但是 60 年代中期兴起的生成语义学才是语言学史上的一次真正革命。虽然不清楚生成语义学家是否对语言的本质属性下过明确的定义，也不清楚是什么原因促使生成语义学 70 年代中期便销声匿迹，但是必须明确提出——语言的本质属性是语义性。可以这么说，注重语形规则系统的研究者始终游离于语言的真实本体之外，而直接研究语义结构规则系统才是语言学

家的任务。人所认识的世界就是语言中的世界,语言的研究就是客观世界在人类精神世界中的存在的研究。即使语义结构研究如同"泥潭",也许不是语言学家能够单独完成的任务,然而如果要把自然语言移植到电脑中去,就必须穿越语义研究的沼泽地。

必须清醒地认识到,耗散性语言符号系统≠递归性数学符号系统,语言使用者≠数理逻辑学家,语言法则≠逻辑规则。语料的非穷举性决定了文法模式的有限性,语义结构的复杂性限定了语言描写的精确度。把语义结构的表层投影的语形作为本体语言学的纯正对象,背离了语言结构的真实本体。语言系统形式化的关键是语义系统的有限形式化,语义标准的主观性和客观性关系本质上是素单位(自然单位)和位单位(功能单位)的关系。乔姆斯基试图在形式句法结构数理化描写的基础上建构生成语法体系注定捉襟见肘,这一初始思路必然导致"抛开语义—深层语义—语义解释—逻辑表达—逻辑式"这样的"不断革命",以致于不可能达到发现语言生成机制的既定目标。依据形式语法的有限范畴和简单框架,既无法驾驭语言结构的生成转换系统,也无法深入语言结构的全面分析,第五代计算机研制的搁浅当在意料之中。

(三)语法的不同平面和研究切口

与西语以形态的自然裸露,诱惑着语言学家围绕着形式语法研究不同,根据构造特征,研究汉语必须从语义入手。从根本上来说,汉语语法研究以语义研究为首选,是由汉语的鲜明个性和语言的本质共性双重决定的。印欧语的形态变化反映着符号配比的语义变化而形成表层"形态语法",因此从形态入手也就曲折地反映着从语义入手。汉语没有反映着符号配比的形态变化,语义选择直接作用于结构而形成"语义语法",因此只要从事

语法研究必然触及语义。任何语言结构都有语义和语形两方面，只是形态语法的重在形态变化而语义语法重在语义结构，语形的手段有所不同。既然语言的构造具备这两个平面，因此从哪一平面入手研究都并非不可，但必须选择合适的研究切口。在操作前务必反复考虑：1.到底哪一平面是人类语言构造的基础，或者最具有世界语言的共性或语言构造的解释力；2.就具体语种而言，最值得入手或最具有解释力的是哪一平面。

就汉语而言，虽然语用对汉语的表达存在着制约性，但语用毕竟以语义结构为基础，或者可以说，语用是一种借助语境和句法结构的选择来表现的语义性。任何一种语言的深入研究，都应当以语义结构研究为基础。正是从这一既是汉语的自身特点又具有世界语言的共性的语义结构入手，语义语法学理论才具有世界语言学的普遍意义。语形型语言的研究法和语义型语言的研究法的世界性合流不可避免，而这一合流的理论基础就是符号序列的语义关联性。

## 四、语义语法学的语域观

(一)语言中的情感法则和事理法则

作为语言赖以形成的人类符号化能力，起源于动物性交换信息的本能，经过人类的早期原始感知而逐步发展出来。早期原始感知并非就是纯理性，而是相当情绪化的驱动。这一特征并未因为人类形成了语言符号系统而销声匿迹，它以底层形式积淀在语言中，不管是语言系统内部还是语言使用过程，都难以摆脱与生俱来的情感。为了在思考和表述中尽量排除情感性和意识流，因此在希腊哲人那里，逐步制定出思辩规则——逻辑——作为辩论的标准。同时形成了符合逻辑规则的语言才是理性语言的观念，由此出现了日常语言、文学语言和学术语言的畛

域。与学术语言的逻辑性不同，无论是经过加工或修饰的文学语言，还是不假思索脱口而出的日常语言，自然语言始终是情绪宣泄的重要途径，因此有“离骚”和“愤怒出诗人”之说。即使在冷静（情绪起伏以后做出的选择）的人际交往中，自然语言仍然包含着不同程度的情绪化色彩。因此语言学的深入研究，不仅要研究自然语言的“反逻辑”法则，而且要研究语言层而下的情感原则。甚至可以说，所谓“语用平面”除了依赖语境的省略，无论是位移凸现还是语气重音都是情感原则的作用。与语义结构的“语义平面”和语形结构的“句法平面”不同，“语用平面”的本质是语言的情感表现。

西方研究的逻辑规则主要存在于思辩语言或学术语言中，只是自然语言法则的部分再抽象。虽然西方语法研究源于希腊的“语言·哲学·逻辑”混沌母体，但是必须分清自然语言语法学和思辩规则逻辑学，因此不取“自然语言逻辑”，而用“语言法则”（简称“语法”）。通常以为词类的语义划分或句子的语义分析就是逻辑分析，甚至还提出与形态和功能标准相区别的概念标准。无论是反对依据“逻辑”给词语分类，还是主张依据“逻辑”给词语分类，都给语法研究仓促套上了“逻辑”的帽子。自然语言单位与思辩逻辑单位不具有严格对应性，实词与概念并不对应。实词“词解”不等于概念“定义”。即使在某些情况下实词语义和概念定义基本对应，但语义仍然大于定义，因为语义还包括色彩义、语体义、情景义和联想义等等。定义必须明确内涵和外延，具有抽象性、精确性和规范性等特征，日常语义不必像定义那么明确，通过指物、比喻、循环等方式就可以了解。正是由于日常语义的意象性、笼统性、游移性、模糊性和个人性，语言才具有如此千姿百态。如果所有的语言表达都遵循思辩逻辑，语言的花朵只可能是令人窒息的黑白。

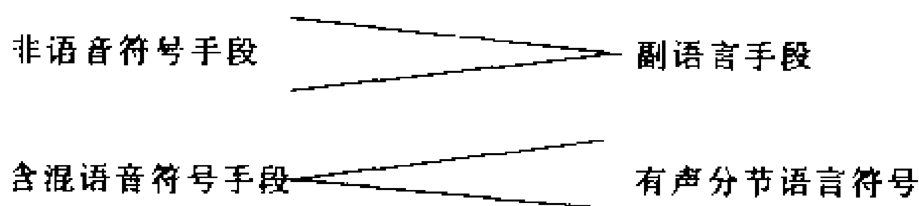
汉语实词的义类划分不是根据逻辑意义，而是借助日常语义；汉语句读之间的语法关系不是根据逻辑关系，而是遵循事理关系。事理关系或事理法则广泛存在于日常语言中，具有时序先后性、空间位置性、主观因果性、主观目的性和评估好恶性等，而逻辑关系存在于学术语言中，具有逻辑推断性和客观因果性等。人们经常为一些事情不平或烦恼，其原因就在于思辩中的逻辑关系和日常中的事理关系之间的冲突，日常事理并非一定依据思辩逻辑运行。事理关系是踢球的随意游戏，逻辑关系是踢球的比赛规则，然而踢球并非定要按照比赛规则进行。不是日常语言法则以思辩逻辑为基础，而是思辩逻辑以日常语言法则为基础，同时在物质世界的研究中抽象出数学符号和几何规则的科学语言。既然思辩逻辑是对日常语言法则或事理法则的提纯和规整，因此依据逻辑关系就不可能对语言法则进行全面描写。语言法则的研究，必须依据事理法则、情感原则、经验框架和文化内涵加以解释，而不是为了体系建构的方便，一下子就跳到思辩逻辑层面。如果思辩逻辑等同于语法规则，那么语法就没有存在的必要。面对日常语言中比比皆是的"反逻辑"，或自发遵循语言法则的习以为常，必须与语言研究中的"逻辑主义"分道扬镳。只有彻底告别了逻辑主义，才能走向生成语义学。

自然语言具有本能传情性和日常认知性，逻辑是对语言法则的窄化抽象和硬性规范，二者虽部分重合但不可混淆。可借用数理符号转写语言法则，但语言法则不是数理逻辑，而是基于认知活动和事理认定并受控于情境原则和情感原则的语言表达法则。语法学家的任务就在于研究自然语法规则，虽然自然语言法则中蕴涵着"自然语言逻辑"，但语法绝不等于逻辑。日常语言是一个耗散性符号系统，数学符号系统是日常计算符号极

度抽象化的产物,是一种封闭的抽象符号系统,因此一些数学家主张用"数学语言"把这一抽象符号系统从自然语言系统中分离出去。

(二)自然语言的层垒包容性

语言能力是我们这些具有意识表现力、艺术想象力的现代人,随着脑子增大在晚近时期从"泛语言"到"现代语言"迅速出现的一种能力。在非语音符号手段和分音节语言之间,有一个漫长的过渡。在这一过程中,分节语言符号日益为主而非语音符号日益为次,呈现出彼消此长的趋势:




人类意识形态从混沌性,经由实物直观性、动作意会性和原始语言性,演变成现代语言性。尽管语言思维成为现代人的本质特征,但是现代人的精神世界中仍然包含着本能反馈、实物直观感受、动作符号意会和原始语言思维,精神世界并非都必须或可能以语言方式表现出来。人类思维交际符号演变的连续统积淀在现代语言中,表现为语言的层垒包容性。除了时间、地域和社会等差异,语言中还存在的层垒性差别,即一种语言的一种地域方言的一种社会方言的内部,仍然存在确实有别但又相互渗透的不同层级,这种层级可称之为"语域"。依据不同的文化素养,不同使用者可能掌握不同的语域语言;根据不同的交际语境,同一使用者在不同场合也可能使用不同的语域语言。语言的"语域"层级大致如图:

| 层级 | 符号性质 | 不同语域 | 作品形态 | 表达特征 | 认知类型 |
| --- | --- | --- | --- | --- | --- |
| 5 | 语言符号 | 思辨语域 | 学术论文 | 逻辑性 | 科学思维 |
| 4 | | 艺术语域 | 文学作品 | 修饰性 | 文学思维 |
| 3 | | 日常语域 | 日常会话 | 明白性 | 日常思维 |
| 2 | 零碎语言符号 | 含混语域 | 零言碎语 | 含混性 | 懵懂意识 |
| 1 | 副语言符号 | 体态语域 | 情态手势 | 意会性 | 动作意会 |

现代语言内部的语域层级系统，与人类语言进化的连续统相一致，人类语言进化的连续性积淀在个人语言的语域层级系统中。以文学作品为形态的艺术语域，当出现在文字产生后；以学术论文为形态的思辨语域，当出现在思辨规则形成中。典型的思辨语域是科学家所具有的语域，典型的艺术语域是文学家所具有的语域。既是科学家又是文学家的个人可能拥有五个语域，文学家则拥有四个语域。非科学家和文学家的个人具有日常语域、非典型的艺术语域和思辨语域。如果是文盲，通常只具有日常语域以下的三个层级，如果发音器官障碍，则剩下了含混语域和体态语域。所谓文学创作就是将感受的结果从日常语域转化为艺术语域的过程，所谓论文写作就是将研究结果从日常语域转化为思辨语域的过程，写作教学的任务就是语域转化能力的训练。

（三）语域语言学的理论价值

语域层级系统与语法规则系统，具有非常密切的关系。迄今为止，语法学家的语形结构研究都是依据一定的学术规范对语言进行解构，而没有理由认为这就是语法规则本身。从规范性出发的语法研究，一方面只是将“理想化”的语言事实尽量描写出来，另一方面又以“不规范”为借口将活生生的语言事实藏

到抽屉底下。由于对通行或书面语近乎偏执的信赖,语法学家对一个文盲所说语言中的法则往往漠不关注。但是如果切实地描写偏僻乡村的农人语言,一定会发现太多的与“标准语法”冲突的现象。然而,农人语言法则并非不是语法,反而是未受书面规范的“本色”语法。20世纪的汉语语法学家,研究对象绝大多数是通行书面语或欧化汉语。有一个不争的事实,越是强调规范性则欧化程度越高,也就距离“本色”语法越远。在语形型语言中,由于词法变化规则明确,凡不符合形态变化规则的语句就诊断为病句。在语义型语言中,以往的语法学家也从“形式”入手,将汉语“病句”诊断为“成分残缺”、“搭配不当”、“词序颠倒”和“句式杂糅”等。如果不包括虚字的调配,其实汉语的句法结构是否合法,完全取决于语义搭配的是否恰当,以上的这些“汉语病句的症状”都是语义问题。

更为严重的是,语法学家经常为之头疼的是归纳出来的语法规则往往没有周遍性,例外层出不穷。之所以语法规则缺乏周遍性,是因为不同语域对表达的要求和语法的规则不尽相同,日常语言要求随机性和明白性,艺术语言要求形象性和修饰性,而思辩语言要求逻辑性和规范性,因此对语法规则的要求也不尽相同。在思辩语域归纳出来的语法规则,在艺术语域可能发现许多例外或所谓“超常搭配”,在日常语域将会遇到更多的例外或所谓的“病句”。书面表达方法和口头表达方法具有本质差别。一方面书语和口语的媒介不同,视觉符号和听觉符号的发送、传递、接受各有特点;另一方面书语除了上下文别无依傍,只有凭借语言符号和个人知识网络去交流,而口语可借助大量非语言符号。正是最大可能地利用非语言符号,才是日常语言中的“病句”层出不穷的内在机制,借助非语言符号可以使用的情境口语,一旦用文字照录不误则成为书语中的“病句”。语言符

号与非语言符号的互补具有反比关系：语言符号的表达越充分，对语境符号的依赖性越弱；语境符号的表达越充分，对语言符号的依赖性越弱。

根据一个托福考了660分的惊人高分的学生的体会，在网上与美国学生聊天时，发现他们发过来的电子邮件，如果用我们的英语语法标准几乎每句都是"病句"。这就显示出国内规范英语教学和国外真实英语交际的差别，一个在中国自以为学会了标准英语的人，竟然觉得美国人"满口病句"，症结就在于所谓"规范语法"僵化，就在于没有清楚地认识到不同的语域具有不同的要求。

面对格里木定律的"凡规律皆有例外"和维尔纳定律的"凡例外皆有规律"，研究语法必须首先进行语言定位，即选择确定的语域、明确语料所属语域以及其抽象规则的适用范围。早年的乔姆斯基万万没有想到，形式语法体系的建立如此步履维艰。生成语法学企图用数学化、逻辑化或形式化手段描写"理想化的语言能力"，但"理想化的语言能力"是什么？而没有明确界定的"语言能力"，可能是一个学术、文学和日常语域混合体的使用能力。理想化的语言能力，只有通过具体的定位才能够落到实处。语义语法学所认定的人类所具有的普遍"语言能力"，首先明确为建立在认知基础上的语义化能力；语义语法学所认定的"理想化的人的语言能力"，首先限定为学者所具备的智性语言能力；语义语法学的研究对象，首先定位于具有逻辑性与规范性的思辨语域。只有在学术语域语法研究完成以后，艺术性的文学语域和生动性的日常语域的语法的研究才有可能展开。乔姆斯基未能明确界定"理想化的语言能力"，注定了形式语言学面对庞大语料势必如坠云雾之中。作为有序异质的语言系统，以相互渗透但规则参差的语域为实际存在，某一语域的特殊规则移置

于另一语域则可能出现缺漏,将不同语域的语言事实和语言规则搅在一起势必治丝益棼。语义结构的法则必须首先限定在某一语域中,只有选准切入口而将较为规则的某一语域的语义形式化以后,才有可能展开其他语域的求异研究。没有对自然语言离析分层,没有给形式语法的研究对象准确定位,同样注定了生成语法学步履维艰的命运。

面对语材的庞杂性,语法的缺漏性,计算机程序的严谨性,能够使研究得以一步步成功的先决关键就是语义语法学的语域观。在语言信息的计算机处理中,理想化的研究程序是:首先研究最规则的学术语域语法;然后以之为参照系,研究文学语域语法,解释与学术语域语法相悖的"超常搭配";最后再以文学语域语法为参照系,研究日常语域语法,在语言符号与非语言符号互补和转化的机制中解释"病句"法则。

## 五、语义语法学的研究体系

(一)语义语法学的研究思路

语义语法学建构的直接原因是自然语言的计算机处理,电脑需要自然语言的语义网络,而只有在这一研究过程之中,才能够更深刻地深入语言的内部结构。

也许,自然语言的处理可以大致划分为文字识别、语言分析、语言生成、语言交际四个阶段。与自然人个体学习语言文字相比,文字识别相当于识字习词,这一阶段已经完成。语言的自动分析相当于自然人的阅读即语言理解,这一阶段目前正在进行,但因为从语形语法学入手而面临重重困难。语言生成相当于自然人的造句即语言表达,这一目标的实现仍然有待时日。无论是自动分析,还是语言生成,实践证明仅靠以语形结构为本位的形式化是无法解决的,传统的词法和句法结构范畴(词类划

分、语形成分、短语结构、句子成分)驾御不了自然语言结构,无法保证语言的正确分析和生成。实现语言分析生成的目标,必须依赖于语义结构的形式化。在此基础上,才有可能研究会话作文、语言翻译等语言交际。

自然语言的计算机处理有不同的术语,根据与“人脑语言学”的对应是“电脑语言学”。从人脑角度出发,电脑语言学的研究是将自然语言移植到电脑程序系统中的电子工程;从电脑角度出发,电脑语言学是电脑程序系统模拟自然语言能力的仿生工程。

不管是系统移植还是能力仿生,首先必须清醒地认识到人脑和电脑的异同。1. 构造机制的不同:人脑是有机生物性神经系统,具有生物认知机制,而电脑是无机电子计算系统,具有电子运算机制。2. 伴随情感的有无:人脑具有作为感觉、思维和表达的内驱作用的情感性,而电脑只具有执行程序与进行运算的机械性。3. 经验基础的有无:人脑具有以直接性、感受性、混沌性和随机性为特征的经验基础,而电脑即使配备了固定的知识背景,也不可能具有经验基础。4. 认知理性的异同:人脑的感受和认知可以区别为非语言层次和语言符号层次,非语言层次包括直觉(内敛的体验)和感觉(外露的体态),语言符号层次包括知觉(游移性印象)、象觉(清晰性意象)和统觉(逻辑性抽象)。电脑的认知理性只能够定位于语言符号的象觉和统觉层次,难以具备丰富的认知层级系统。5. 交流对象的异同:人脑的交流对象分别为对象世界和人际之间,具有交流的二重性;而电脑的交流对象只有执行可接受程序的人,仅具人机对话的一重性。计算机不可能直接认知世界,互联网交流仍然是执行者之间的交流。

其次,必须明确地认识到人脑语法与电脑语法的不同。人

脑语法或自然语法是在历史上逐步形成的,具有规约性、类推性、随机性和耗散性特点;而电脑语法只是自然语法的再抽象化和再规范化,具有规范性、是否性、程序性和封闭性的特点。适合于计算机处理的电脑语法不等于人脑语法,电脑所能获得的人脑语言能力只是有限语言能力。移植进电脑的自然语言必须进行界定:1. 定域化的语言:因为人脑语言是一个可以分为不同语域的复合性层级系统,所以必须首先确定电脑语言与人脑语言的对应层级,最佳选择就是学术语域。2. 形式化的语言:其次必须将所定语域的语义和语形分别形式化。3. 数字化的语言:再次必须将形式化的语义和语形分别数字化。4. 编程化的语言:最终将数字化的语义和语形进行计算机语言的编程。语言的计算机处理的目标,就是在象觉性与统觉性的语言认知层面上实现人机对话。归根结底,电脑语言只是一种依据电脑性能而精心设计的规范化自然语言,或者可以说是一种特殊的人工语言系统。电脑的运算速度远远超过人脑的思维速度,但是电脑不具备能动的认知行为,与客观世界的交流必须以人的认知能力为中介,因此电脑智慧可能胜过个人的智慧,但永远无法超越人类的智慧。

为了实现自然语言法则的形式化,其关键就是必须实现语义网络的形式化。乔姆斯基的生成语法学,只是在现代英语所承袭的欧洲先民的认知语义范畴形式的描写基础上,企图建构现代语言的形式句法大厦,并以揭示人类的普遍语言能力,但是由于不愿放弃形式句法规则的基础地位,因为放弃这一基础地位即宣告语形语法学的终结,因而陷人自身不断“革命”的颠沛流离。

语义语法学的研究目标就是实现语义网络的形式化,但是语言学家往往视语义结构为非语法,视语义结构研究为畏途,

“泥潭”情结误导人们认为语义结构的形式化不可思议。一方面语义单元不如语形标志那么裸露，语形标志具有可视性标准，而语义单元含有较多的非可视性意会；另一方面语义单元并非没有范畴性标准，语义单元具有感受或认知的一致性。一方面语义范畴不像语法范畴那么有限，语义系统不像语形系统那样封闭；另一方面尽管语义单元数量众多，语义结构关系复杂，但它们具有潜在的元语言系统。

语义网络由语义单元（符号实体）和语义结构（符号关系）组成。语义单元具有聚合关系，而语义结构则是组合关系。就语义单元而言，所谓语义的主观性和客观性也就是语义的个人联想性和集体认同性，其本质就是自然存在的“素单位”和社会存在的“位单位”之间的矛盾。语义语法学的对象是集体认同性的“义位”（“义素”的概括），而不是个人主观性的“义素”（“义位”的变体）。就语义结构而言，语义结构的形式化就是句法结构的形式化，语义结构形式化是语形结构形式化的实相，而语义结构形式化的实质就是相关语义场的可关联性的形式化。因此，语义网络的形式化包括语义单元的形式化和语义关联的形式化两个方面。每一语义单元在语义场中具有两种不同类别的语义，一种是同一语义场内的各个“义位”之间的“义征”异同，另一种是不同语义场之间是否具有语义结构关联性的“义联”异同。针对前者，采取解释性元语言进行语义分析而建立聚合性的“义位”网络系统，针对后者，采取关联性元语言进行语义分析而建立组合性的“义联”网络系统。

（二）语义语法学的研究方法

语义结构的形式化方法是语义语法学的基本研究方法，语义结构形式化的预设是自然语言系统中存在一个与基本认知能力密切相关的元语言系统。这一元语言系统，可以分为解释性

元语言和关联性元语言。与乔姆斯基着眼于语形结构而采取演绎法不同,以语义为本位的语法研究立足于有限网络模式下的归纳法。所谓有限网络模式,是对耗散的语义网络系统所做的简化。

1.义征标记法

"义征"即语义特征,不取通常使用、然而不当的"义素"术语。依据有限网络模式,借助义征分析对语义单元进行形式化描写或义征标记。要进行义征标记,务必建构义征系统或解释性元语言系统。具体的途径是,首先归纳出汉语普通话中可满足解释确定的词汇系统的基本解释性词语(可能在1500个左右),然后对基本解释性词语加以删汰、合并和分类,抽象出义征范畴以建立解释性元语言的层级体系。与通行义征分析不同,不是为分析意义而分析语义。以解释性元语言作为标准的义征分析法,其目标是建立以解释性语义场为单元的语义聚合网络。

2.义联标记法

"义联"即语义结构关联,不取通常术语"语义选择"或"语义制约"。依据有限网络模式,进行语义关联的细致剖析,对语义符号配比进行形式化描写或义联标记。语义配比的实体是关联结构体,包括NN关联结构体、NV/NVV关联结构体、AN关联结构体、AV/VA关联结构体等,要在分析这些结构体的过程中归纳出关联性元语言或义联标记。语义结构并非个别符号之间的搭配,而是语义场之间的配比关联,因此义联标记法中包含了义场配合的原则。在具体研究过程中,以最小语义场为基础,归纳出相关语义场可能的关联要素。以关联性元语言作为标准的义联分析法,其目标是建立以关联性语义场为单元的语义组合网络。

目前已有的配价语法和格语法的语义结构描写方法和成果

可以参照，但是语义语法学要求语义语法的各层子系统的描写目标服从于建立句模系统的总目标，务必根据汉语的特点来确定关联性元语言系统。配价语法的立足点是动词核心论，但汉语中存在非动核性句干，仅靠"动词核心论"分析模式不能全面描写汉语句子结构。

3.语形标记法

汉语没有形态变化，但是具有词序、虚词和语调三种形式手段，其中虚词是具有特色的一种语形系统。汉语的语形包括词组级语形和句子级语形，前者可以称之为完构成分，后者可以称之为完句成分。完构成分附着在词上，如结构助词、时体助词等；完句成分附着在句干上，如语气助词等。依据语形范畴系统和语形标记法对语形系统进行层级性形式化描写。

4.标记合成法

在义征标记、义联标记和语形标记的基础上，通过相关标记的合成，可以建构句模系统，完成汉语语义语法系统的描写。

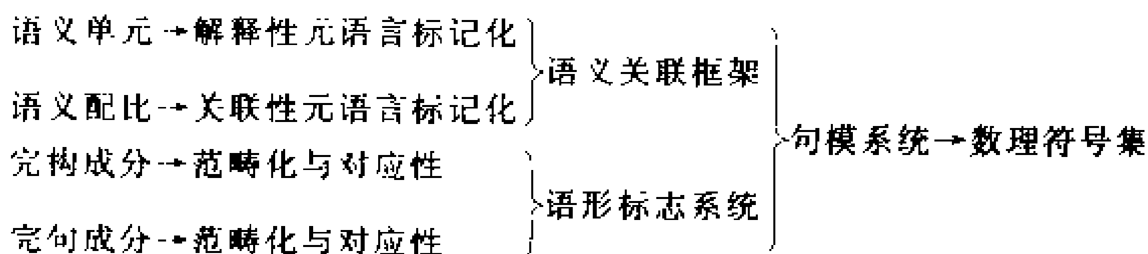
5.计算验证法

汉语语义语法系统研究的每一步结果，都可以也应当用计算机操作系统作为验证手段。

（三）语义语法学的研究步骤

虽然汉语没有西方意义上的语形，但不等于汉语的语义结构不能形式化。现代汉语语言系统形式化的步骤大致是：1.用解释性元语言描写语义单元的"义位"，给出语义场之内的语义联想或区别特征，组成层级语义系列或聚合性语义场。2.用关联性元语言给出语义配比的"义联"，给出语义场之间的语义选择规则或语义制约，构成生成符合自然人认知的语义结构体的语义关联框架或组合性语义场。在语义单元聚合范畴化和语义结构配比范畴化的基础上建构语义网络。3.进行完构成分和完

句成分的范畴化研究,寻找语义关联和语形成分之间的对应性,给出语形标志系统。4.以语义关联框架和完句成分系统为基础,建构汉语句模及其系统。5.将句模系统转写为数理符号集。汉语语义结构逐层形式化的步骤如下:




尽管语义的形式化相当困难,但是在语义语法学理论的指导下,根据定域(确定自然语言的语域)、定量(确定语义单元的数量)、定性(以元语言确定语义单元的语义关联的性质)、定位(确定语义单元在语义网络中的位置)的"四定"原则,建立适合本思路的定域性语料库以逐层实施,自然语言的形式化至少可能在某一语域的单句模式的系统中首先实现,然后才有可能进一步研究这一语域的叙述结构(复句、句群、篇章)的形式化。

(四)语义语法学的研究目标

语义语法学具有理论研究和应用研究的双重目的,其理论价值在于建构符合人类语言共性和语法真实本体的语法学理论,其应用价值在于促进语义网络的形式化研究,为语言信息的计算机处理提供可编程文本。语义结构体系的形式化研究可以形象地称之为"语言基因图谱分析工程",元语言好比语言基因。与拘泥于"窄化语法"而难以跳出"语形框架"不同,语义语法学抓住语法的真实本体,通过三万个词的解释性义征和关联性义联的研究,可以完成语言基因图谱分析工程。在语义语法学引导下,以认知语义和表达结构的研究为突破口可以贯通认知语

言学、心理语言学、功能语言学、文化语言学、社会语言学、数理语言学和计算语言学等。

以往的语法学研究，除了执著于语义结构的表层投影以外，还有一个根本缺憾，由于不具备自然科学研究那样的实验手段，因此难免经验性、臆断性和游移性。语言基因图谱分析工程将采取计算机作为验证手段，使语言法则的研究具有自然科学性和可证伪性，在语义本位的基础上促成21世纪的语言科技革命。语言基因图谱分析工程具有科学性、创新性、可操作性、可验证性和可扩展性，研究成果具有预期性和技术化前景。总而言之，语义语法学的前途就是将人脑语言机制移植到电脑中去，从而实现电脑对自然语言的理解和生成。

但是事情并没有全部结束，还有任务在等待着专家继续探索：

1. 如果有必要，进一步进行文学语域和日常语域的语义网络形式化研究，促使电脑人在学术语域能力的基础上兼具文学家和平常人的语域系统。尽管在继续的研究中仍然会遇到新的难题，但是既然有了学术语域系统做参照系，文学语域和日常语域的形式化并非不可能实现。

2. 电脑人对自然语言的深入而扩展的理解还需要文化知识网络。获得了自然语言能力的电脑人，还需要继续上中学、大学，攻读各门学科或专业的硕士和博士学位。掌握了文化知识网络的电脑人，将被塑造为一个百科全书式的通才。

3. 从语言的使用过程来看，语言除了语义和语形编码，还具有语用编码。语言的功能是表情达义，因此广义语法的研究不可避免地要涉及语用编码或语境手段。电脑人只有掌握了语言符号与语境符号的转化原则，说出来的话语才能像自然人的话语一样得体。

4. 人类语言除了明智能力还具有情感能力，从语义的“泥

潭"中爬出来以后,还不得不面对隐藏在语言规则之下然而参与语言表达的情感规则。人类情感机制的形式化是心理语言学家的任务,只有对电脑人进行情感或心理机制塑造,电脑人的表达才不至于毫无人情味。

因此,广义形式语言学的体系可以图示如下:

| 认知性语言符号的形式化系统 | |
| --- | --- |
| 语形结构的形式化系统 | 语用结构的形式化系统 |
| 语义网络的形式化系统 | |
| 文化知识网络系统 | |
| 心理情感机制 | |

毋庸置疑,语义网络的形式化将大大深化人们对语言自身和语言法则的深入了解。只有在语义实体、语义结构和语形系统详尽描写的基础上,认知语言学才可能真正揭示出语言结构和世界图式之间的像似性秘密。

只要探索语言本质的好奇永远存在,只要人脑语言移植电脑的目标不会改变,就必须穿越语义结构的沼泽地,而可能的引导就是在前修时贤的基础上,建构以人类语言共性为基础的语义语法学(semantic grammar)。语言学家务必彻底消除"语义泥潭"的畏缩情结,语义结构的研究,不但是语言学家而且也仅仅是语言学家才能承担的根本任务。语义语法学的创建,既为国际语言学和汉语语法学提供了崭新的理论,又为语言信息科技工程找到了可行的思路。19世纪是历史语言学的世纪,20世纪是语形语言学的世纪,而21世纪将是语义语言学的世纪。

# 第七章　东西方文化的本原差别及其分野

从距今3.5万年前开始，心智完全成熟的现代人类所创造的一切都可以称之为“文化”。尽管考古学界的一些学者为了区别不同的阶段，把史前现代人的创造称为“文化”，而把史后现代人的创造称为“文明”，但通常习惯于用“东西方文化”来指有史可考的欧亚非文明。东西方文化的差别主要是指史后文明的差异，但在讨论文明起源时必然牵涉到史前文化，因此东西方文明也就包含了相关的东西方史前文化。

季羡林先生提出：1.“经多少年思考，世界文化可分为中国文化、印度文化、伊斯兰文化与西欧文化”；2.“前三种文化为东方文化，西欧文化为西方文化”；3.“西欧文化为印欧语言”。[①]虽未见详细论证，但季先生对东西方文化的多年思考和东西方文化类型或世界文化类型的划分，却给文化学提出了一系列问题：1.“东西方文化”的观察点是什么？2. 世界文化类型划分的标准是什么？3. 这四种文化的文明源头是什么？4. 语言文字谱系、语言文字类型与文化类型是什么关系？5. 语言类型、语言分析模式与思维模式和本原哲学模式有着什么样的关系？6. 科技文化、艺术文化与文化类型有着什么样的关系？7.“东西方文化”的合适类型划分应是什么样的？

---

① 季羡林《神州文化集成·序》，新华出版社1991年。

这一系列问题不仅是区分东西方文化类型或世界文化类型的关键,而且关系到一般文化哲学的基本理论和对整个世界文化历史和类型的宏观把握与中微观研究。迄今为止,尚未见到这方面的追根穷源、纵横比较的系统探讨,而尝试展开大文化的本原性、跨学科、全方位探讨,对"读懂世界或理解中国"可能有所促进。

## 一、现代地理学分界与文化中心主义的视角

作为近现代地理学概念的亚洲和欧洲,并不是文化学上的文化类型或文化圈的概念。把处于亚洲的文化称之为"东方文化",把处于欧洲的文化称之为"西方文化",以近现代地理学的概念作为区分东西文化的标准,虽然出于某种习惯,但显然缺少文化学理论根据,至少可以说是囿于近现代地理学眼光。

东西方的地理概念,不仅具有空间上的相对性,更重要的是具有时间上的延展性。以欧洲为观察点,习惯上将位于欧洲之东的亚洲分为近东、中东和远东。如果以从古到今的欧洲为观察点,"东方"则是一个逐步向东延伸的动态概念。古希腊以小亚细亚为东方,古罗马以土耳其为东方,中世纪以阿拉伯为东方,殖民时代开始以印度等地为东方,而后以中国等地为远东。将地处欧洲之东的亚洲的不同种类的文化一概称之为"东方文化",无疑是出于欧洲学人长期以来形成的以"西欧中心主义"为观察点的"西欧文化眼光"。

以现代中国为观察点,印度地处南亚,伊朗等地处中亚,阿拉伯区域地处西亚,欧洲可统称为西欧。如果以从古到今的中国为观察点,"西方"也是一个逐步西移的动态历史概念。上古三代以中原为大地之中央,商代始称"商中",周代称"国中"、"原中",后演变为"中国"、"中原"。所谓周边四方民族分别称之为

东夷、西羌、南蛮、北狄。源于西羌的周族来自殷商的"西方"今陕甘青一带。汉代称今中国境内的新疆，包括中亚、西亚一带为"西域"。唐代称印度为"西天"。明初称东南亚、南亚、西亚，甚至包括东非一带地域和海域为"西洋"。明末满清以"西洋"称欧洲，亦称之为"西方"、"泰西"和"大西"。与西域、西天相对，汉唐帝国的中原又自称为"中土"、"东土"。因此秦汉以后，以中原为观察点，中原以西则谓之西方。依据华夏汉人这一历史性视角，在中原汉文化区域之外的西部文化皆谓之西方文化。此出于以"东亚中原主义"或"中国文化中心"为观察点的"东土文化眼光"。

在历史上，如果以阿拉伯与欧洲相对，则阿拉伯自称为"东方"。阿拉伯哲学家伊本·西那(980—1037)的一本著作就叫《东方哲学》，以致于"东方学"术语在西方沿用至今。如果以印度为观察点，则东西方又各有不同。以地理位置的相对性指称传统的"东西方文化"，既基于地域上的东西之别，又隐含着"西欧眼光"与"中原眼光"的不同的"文化中心主义"，并不是不同类型文化的本质划分。

显而易见，东西方文化类型的划分标准不应是欧亚两洲的地理学分界，不应是西欧中心主义或东亚中原主义的文化视角。世界文化类型的划分，应当运用以源头文化、语言结构、本原哲学和衍生文化四维组合而成的综合性标准。

## 二、地中海文化圈和东亚南洋文化圈

远古地中海文化圈包括苏美尔—巴比伦文化、古埃及文化与爱琴海文化(基克拉泽斯—克利特—迈锡尼)。依据文明的起源，伊斯兰文化(西亚—阿拉伯—伊斯兰)、西欧文化(古希腊—罗马—欧洲)和印度文化(古雅利安—吠陀—佛教)的母文化源

头皆在地中海文化圈。尽管苏美尔人来历不明，然而苏美尔文明无疑是西亚文化的古老源头，同时古埃及文化是西亚文化的接触源头，犹太文化、巴比伦文化及亚述文化为之后裔。阿拉伯半岛插入埃及与巴比伦两大古代文明之间，南部海路又与古代印度文明相通，很早就接受这些文明影响的阿拉伯人，在公元6世纪形成伊斯兰文化。

爱琴海文化与两河流域文化有惊人的相似性，又与埃及文化早有接触。迈锡尼人一度深入腓尼基、比布鲁斯和巴勒斯坦，形成塞浦路斯—迈锡尼文明。多利亚人的古希腊文化源于爱琴海文化，又受到西亚文化的影响。文献证明，希腊神谱与巴比伦神话极其相似，而巴比伦神话则源于苏美尔文明[1]。前9世纪，希腊从腓尼基借来字母，形成了古希腊文明，罗马继承和发展了希腊文化，由此演变为西欧中世纪文化、文艺复兴文化和近现代科技文化。欧洲的整个精神生活，包括哲学、道德、美学和政治思想，都源于希腊—罗马的思想家和政治家的精神创造。

印度远古文明是由达罗毗荼土著创造的哈拉巴文化，大概在前1750年开始衰落。前1300年，一支来自中亚的与现代欧洲人同根，所操梵语与欧洲语言同源的雅利安人进入印度。在从牧畜民转变为半农半牧民的过程中，创造了恒河吠陀文化，在前6世进入佛陀时代。雅利安创造的印度文明源于地中海文化圈，虽分化较早但一直交往，婆罗米字母即源于西亚阿拉马字母。前6世纪和前4世纪，大流士一世和亚历山大大帝曾先后占领印度的部分领土和入侵印度河流域，把波斯文化和希腊文化传入印度。

---

① ［法］皮埃尔·韦尔南(J.P.Vernan)《希腊思想的起源》(秦海鹰译)，98页，三联书店1966年。

与西欧文化、印度文化、伊斯兰文化都是地中海文化的变元文化不同，华夏文化植根于远古东亚文化圈。依据考古文化学研究，大约距今7000年前后，东亚大陆形成了东南青莲岗、西北仰韶和北方细石器三大考古文化系统。距今6000年左右，在中原地区形成了一个兼包东西考古文化系统特征的文化体，这一庙底沟氏族部落文化可能就是原始华夏文化共同体的胚胎①。三大考古文化系统交汇的中原地区，经过持久的接触、冲突、交融和洪溢，在距今5000～4000年左右形成了原始华夏文化，历五帝三代日臻成熟，以汉唐文化蜚声世界。虽然根据四川三星堆出土的，不同于中原文化的象征王权的金杖、覆以金面罩的雕像和青铜神树等文物可以证明，距今4000～3000年左右，西亚文化曾经传播到长江中游的巴蜀地区②。但是，这并不影响华夏文化是一种单独起源，在发展过程中受到西来文明影响的文化的论断。

由于生态环境的差别，就主要经济方式来说，远古地中海文化圈的主流文明是游牧—通商型文明，远古东亚文化圈的主流文明是农耕型文明。

## 三、西方神质宗教文化和中国人伦宗法文化

在远古时代，不同的自源文明都自发产生过原始自然神崇拜与祖先神崇拜，由此构成了人类的神话或童话时代。随着生态环境的对文明发展的制约，东西方对神的崇拜则有所侧重。内陆海的地中海适合海上贸易，为南岸的古埃及文明、东岸的亚细亚文明、北部的爱琴海文明提供了文化交往的途径。远古时

① 详见石兴邦《中国新石器时代考古文化体系及其问题》。
② 屈小强等《三星堆文化》，四川人民出版社1993年。

代地中海上的商业活动，孕育出蓝色的海洋文化。东亚大陆的北部是草原和戈壁，西部是大漠和高山，东南部是丘陵和湖泽。虽然东部沿海地区曾经产生过与通商型海洋文明不同的捕鱼型沿海文明，甚至有人提出古代中国、朝鲜和日本之间的海域为“东亚地中海”之说，但是在中国北部海洋那边的、远古时期从大陆或南亚传过去的朝鲜和日本文明都不是原始的本土文明。有证据表明，远在东亚大陆南部的南岛文明也源于东亚大陆文明。也许在远古时期（距今 2 - 1 万年），由于蒙古利亚人的迁徙，曾经存在过一个“泛太平洋文化圈”，虽然南岛居民被称之为“善于航海的居民”，在岛屿之间有着海上交通的路线，但面对汹涌的太平洋，没有地中海上的那种往返航行的风平浪静，很难以海上贸易为媒介促使这一广袤的海洋圈内多种文明的交融。因此与地中海文化圈不同，古代东亚大陆三大考古文化系统的交汇之处在河洛地区，由此整合成黄色的内陆文明。

居住在畜牧—通商民型的地中海文化圈的人们，也许由于濒临海洋，面对海洋的力量引起遐想和探索，从而演变为以自然神为主导的崇拜。生活在农耕民型的东亚文化中的人们，也许由于植根黄色的土地，面对土地的万物产生顶礼膜拜，从而过渡到以祖先神为主导的崇拜。是崇拜自然神为主还是祭祀祖先神为主，是倾向于呼唤个人内省还是倾向于强化家族外抑，东西方从远古精神信仰中孕育出神学信仰宗教与人伦礼教宗法的两大传统。依据精神信仰的主要倾向，西欧文化、印度文化和伊斯兰文化是源于自然崇拜的、思考神质的内省型宗教文化，华夏文化是源于祭祀祖先的、倡导人伦的外抑型宗法文化。

思考神质的宗教文化或称之为神学文化。远古时期两河流域的居民崇拜自然力，在苏美尔—阿卡德时期，代表自然力的神变为了城市的保护神，至巴比伦时期形成马都克创世神话。同

时两河流域也流行家族制思想的祖先崇拜和信奉天象征人世吉凶的星宿崇拜。西方宗教理论的创立，源于苏美尔文化（前第三千世纪）留下的，以自然神崇拜为主的对于创世神质、人间善恶和个人命运所具有的理性思考。如果把苏美尔人的创世神质看作自源宗教，那么在此基础上孕育出的犹太教（前13世纪）则为直接次生宗教，基督教（公元前后）与伊斯兰教（601）都是间接再次生宗教，其共同特征就是“以自然为最高神祇，把智者变成尘世教主”。基督教源于犹太教，以犹太教的《圣经》为《旧约》。西亚经由犹太文化、巴比伦—亚述文化、波斯—伊朗文化，至610年在阿拉伯世界出现了以《古兰经》为圣典的伊斯兰教，但并不排斥“摩西”和“耶稣”在宗教上的同等地位，形成了伊斯兰文化。犹太教和基督教对阿拉伯人的影响可以追溯到公元1世纪，由于二者带有明显的异族入侵的文化背景，有悖于阿拉伯民族感情，才产生了阿拉伯人自己的宗教，但其宗教原理为借源次生。印度的婆罗门教（前10世纪）奉“梵”为最高神，认为梵天为世界本质，万物皆由此生，以《吠陀》为经典，其自然神崇拜的痕迹十分明显。创立此教的雅利安人由中亚一带迁人，其宗教原理有可能源自两河流域的早期理性主义，也许还接受了哈拉巴文化遗留下来的生殖崇拜影响。继起的佛教（前5世纪）原理源于婆罗门教。与基督教相比，佛教的特点是以缘起论排除了神创说，可以说基督教是“神学宗教”，而佛教是“哲理宗教”。与伊斯兰教相比，佛教以“先知说”排斥了“教主说”，可以说伊斯兰教是超验性的“真主宗教”，而佛教是宣扬四谛的“觉悟宗教”。

宗法文化或曰人伦文化、礼教文化、法治文化。先秦的道家、儒家皆是人伦哲学或伦理哲学，而不是宗教神学。老子取象于人之出生之“道”（产道），比类万物产生之“道”（天道），所谓“天人合一”，以效法生育之“无为”（母子化育顺其自然）为君王

南面之术（无治而无所不治），实为源于生殖崇拜的“化育哲学”。与道家具有渊源关系的儒家，从“母子化育”延展为“仁者爱人”。孔子取象于“仁”、主张“正名”以“克己复礼”，实为纹饰化的“礼教哲学”。老子重在研究自然人或人的自然性，孔子重在研究社会人或人的社会性。老子主张以制欲为手段返璞归真，孔子主张以伦理为纲纪克己复礼，其深层一致性在于以血亲关系或家族伦理为本位，因此可统称为“人伦哲学”。由此又演变为荀子—韩非子以“法、术、势”为核心的法家“权术哲学”，以严刑峻法为手段，形成绵延几千年的“名儒实法”的专制传统。西汉以降，以道家为缘起、掺以巫术而衍生的道教，目标是以“气”养“性”而得道成仙，与西方通过忏悔以追求心灵平静的“原罪”宗教不同。汉魏以降，佛教西来，但在中土始终没有成为信仰的“觉悟”（“佛”之义“觉悟者”）宗教。除了历史上多次发生的灭佛事件，在民间流行“谤佛讥僧”之风。因此我在多年前就提出：中国人有恐惧而少信仰，有潜在的“恐惧情结”而很少升华到自觉的“信仰”。在宗法制度文化之下，从集体性恐惧原型中产生的只是祈佛保佑、趋吉避凶的迷信和安分守己、自解自叹的命运。

汉语的“宗教”其实质为“宗法”、为“伦理”、为“礼教”。《说文》：“宗，尊祖庙也；从宀示”，段注：“示谓神也，宀谓屋也”。子孙祭祀祖先神人于祖庙之义。《说文》：“教，上所施，下所效也；从攴孝”。“攴”，小击也，父执棍教子之形。“教”与“孝”、“效”同源。所谓“宗”与“教”就是祭祀祖先以上行下效，维护家族统治的主要仪式。在这种主流文化背景下的中国社会，其本质只能是宗法主义。就西欧的“宗教”词源来说，姑且举英语 religion 为例，其义是：1. 信仰；2. 一心追求的目标；3. 有关良心的事情；4. 自己感到应做的事情；5. 认真严谨的。综合本义“信仰”和其余转义，religion 的含义：“从良知出发、认真严谨地做自己感

到应做的事，以忠实于一心追求的目标即信仰”。并非中世纪的欧洲没有教会的恐怖，但那是后起文化现象，这里主要就本原文化而言。以汉语词“宗教”翻译 religion 并不妥当，合适的翻译可能是“信仰”。

中国传统文化中没有 religion，只有宗法。与西方 religion 的个人信仰内省性特征相比较，中国“宗法”的家族伦理强制性特征非常明显。在传统宗法主义或礼教文化中，缺乏发自个人内心的虔诚信仰与价值追求，只有处于家族统治下的外在恐惧与无奈祈祷。在中国集体无意识的原型中，没有主动的净化灵魂的向善为美的仁慈和博爱，只有被动的明哲保身的趋吉避凶的侥幸和冷漠。西欧宗教文化中，人可以作为“个人”活着，自愿与上帝（良知的神圣化）交流。中国宗法文化中，人只能作为“类人”活着，被迫受礼教（节欲的规范化）束缚。这不但是基督教文明与儒家文化的根本差别，也是东西方传统的生命价值（个人价值优先和集体价值优先）、组织模式（集权制与分权制）和组织规范（重在惩治民众的律治与重在限制权力的法制）及其社会心理结构的本原性差别。

## 四、西方屈折语言文化和中国词根语言文化

人类文明和文字最早的创造者——苏美尔人其语言谱系归属不明，但依据苏美尔文字的研究，其语言结构属于粘着和造格，可存而不论。根据历史比较语言学的研究，埃及语言、巴比伦—亚述语言、阿拉伯语言属于闪含语系，因此埃及文化、巴比伦—亚述文化和阿拉伯文化都是闪含语言文化。同样，印度语言、希腊语言、罗马语言属于印欧语系，皆源于原始印欧语，因此印度文化和希腊—罗马文化同属于印欧语言文化。闪含语言和印欧语言都是屈折类型，也许有一个更为远古的共同来源，换而

言之,闪含语言文化和印欧语言文化可能源于同一原始母文化。闪含语言是辅音词根的屈折型,一般由三个辅音构成词根,通过辅音之间的元音交替变化以构成不同的语法形态和词性转变。印欧语言是元辅音屈折型,既有元音屈折,又有词尾附加。也许,闪含语言比印欧语言保存着更为古老的原始状态。

组成"东亚南洋语言文化圈"的南岛—南亚语、藏缅语与阿尔泰语的原始结构都是粘着型。正像原始华夏文化是混合文化一样,原始华夏语是混合语,与毗邻三大语言先后有过不同历史阶段的层叠性发生学关系。经过多次的持久融合,华夏汉语由原始粘着型演变为词根语或孤立结构,主要采取语序、虚字和语气的手段来表示语法意义。上古汉语以复辅音的单音节词为主,也许还有辅音后缀,即 CCVCC 结构。

如果依据语言的结构类型,那么西欧文化、伊斯兰文化、印度文化是屈折型语言文化,中国文化是孤立型语言文化。作为研究印欧吐火罗语的语言学家季羡林先生,在论及"西欧文化是印欧语言"的同时,却把属于印欧语的印度文化划分给与印欧语不同的东方文化。由于自觉不自觉以"西欧眼光"来观察"非西欧文化",因而忽视了这一显而易见的语言亲属关系,未能认识到语言结构类型在文化结构类型划分中的重大价值。

## 五、西方音素论文化和中国音节论文化

有了文字,人类才跨进了文明史阶段。苏美尔人创造了楔形文字,由阿卡德人传至巴比伦—亚述人,最终由波斯人改造为音节文字,这就是历史上的楔形文字文化圈。古埃及人创造了圣书字和分析语音的离素原则(音节结构分析到音素),闪米特人在借用圣书文字的基础上创造了辅音字母表(语音系统分析到音素)。闪米特字母西传至希腊,希腊人改造成为元—辅音字

母表(所具有的音素或音位都由字母表示),再通过拉丁文化流播于欧洲。闪米特字母东传,印度人改造为音节字母,再衍生出南亚和东南亚多种字母。闪米特字母在阿拉伯地区演变为阿拉伯字母,通行于伊斯兰世界。希腊字母、拉丁字母、斯拉夫字母、婆罗米—梵文字母和阿拉伯字母都源于闪米特字母,在历史上形成"闪米特字母家族"。

中国的文字最早为殷商甲骨文,迄今为止,尚没有证据表明甲骨文为外来文字。与西方文字的拼音化方向完全不同,由于词根语结构类型的制约,汉字走上了由假借原则(由同源通用,到同音借用)到形声原则(由添加意符或声符,到意符和声符拼合)的发展道路。汉字除了在中国境内为周边民族借用和仿制之外,还东传朝鲜、日本,南传越南,在历史上形成"汉字文化圈"或"汉字家族"。

依据文字的发生学分类,可以分为自源和借源文字。欧洲、印度和伊斯兰文化的文字都是借源文字,其源头文字早已死亡。汉字不仅是自源文字,而且是至今还在使用的惟一的自源文字。汉字是汉字谱系的本元文字,欧洲、印度和阿拉伯文字都是闪米特文字谱系的变元文字。语言的类型对文字的类型具有制约性,东西方不同的语言类型决定了东西方文字类型的不同发展方向。依据文字的功能分类,可以分为表音类型与表意类型。欧洲、印度和伊斯兰文化是表音类型文化,中国文化是表意类型文化。

在文字类型深层,蕴藏着西方的"音素论"思维模式和中国的"音节论"思维模式。音素论源于音素组合的离散继起性,音节论源于音素组合的混成共起性。所谓离散继起性,是指音素结合的界限若连似断,听感上线性流逝;所谓混成共起性,就是

指音素结合的界限紧密难分,听感上瞬间一体[1]。暂举英语和汉语为例,英语的sway(摇摆)与汉语的suei(岁),国际音标同为[suei],但在实际发音过程中,sway的发音表现为s-u-e-i,音素之间的界限较明,时值较长;"岁"的发音表现为u{s/ei},即先作合口状,然后将声调包裹着的整个音节脱口而出,混然一体,时值较短。屈折语的word(汉译为"词")[2]常为多音节,多音节词中的单音节没有语义,必须几个音节组合才能成为word。这种处于待拼状态的纯表音音节,具有明显的离散性。此外,元音屈折和辅音词缀进一步强化了屈折语音素的离散性。反之,词根语中的"名"是完整的音义体,一个音节表现一个语义。汉语音节的附义性,进一步强化了音素的共起性和音节的混成性。

由屈折结构类型的音素继起性自发产生了闪含—印欧语语音分析的音素论。音素论认为音素是构成语音的最小单位,通过音素的组合构成音节,由音节再构成词音,词语的形态变化和词性转类以音素的交替和附加为手段。音素论具有分析性、因子性、组合性的显著特点。由词根结构类型的音素共起性自发产生了华夏汉语语音分析的音节论。音节论认为音节是语音的自然感知单位,音义二者同源,音节无须组合就可以表达完整的词义,语词没有形态变化而词性转类通过音节的孳生为途径。音节论具有意象性、整体性、孳生性的显著特点。依据自发的语音结构分析模式,欧洲、印度和伊斯兰文化是音素论思维模式文

---

① 关于继起性和共起性的概念,参见余志鸿《语言的继起性和共起性原理》,《文字与文化丛书(二)》,光明日报出版社1887年。余文注:继起性和共起性的概念引自桥本万太郎。

② 现代汉语的所谓"词"古代叫"名",与word性质不同。如同orange在英国指"橙子",汉译为不同于orange的"橘子"。

化，中国文化是音节论思维模式文化。

导致人类文明差别的缘起，除了生态环境及其经济方式，最关键的就是具有沟通自然和精神的环节——语言模式。前第三和第二千纪，埃及王国产生的音素论模式和商周王朝产生的音节论模式，才是东西方文化本原性差别之缘起。

## 六、西方思辨逻辑文化和中国顿悟意象文化

美国语言学家比克顿（Derrick Bickerton）在《语言和物种》（Language and Species）一书中很有说服力地表述了这一点："只有语言能够冲破锁住一切其他生物的直接经验的牢笼，把我们解放出来，获得无限的空间和时间的自由"①。毫无疑问，人类学上的心智现代人类是语言符号的动物，由于人的存在，物质性的世界才具有了的意识性的语言。

人所认识的世界就是语言符号中的世界，语言符号的建构方式成为语主认知系统中的世界建构方式。所谓"人类的思维具有共同性"是指大脑功能的共同性，而不是指思维模式的一致性。反之，不同的语言模式具有不同的认知或思维模式。认知模式偏重于动态探索过程，而思维模式偏重于静态思考程序，为了方便可以统称为思维模式，语言模式与思维模式具有相关性或同构性。具有层级性的思维模式，可以分为日常会话模式和思辨论证模式。思辨论证模式是一种发现模式，在西方表现为探索自然的科学模式，在中国表现为阐释道理的名辨模式。认为某人的话语"不合逻辑"，是指不符合源自西方的形式逻辑，但并不意味着在日常生活中大量存在的这些"不合逻辑"的思维或话语没有"自然语言逻辑"。因此在形式逻辑之外定有日常思维

① 转引自理查德·利基《人类的起源》，92页。

或自然语言的逻辑，只是我们受西方形式逻辑影响太深，将一些自然语言法则斥之为"不合逻辑"。经常抱怨或谴责生活中的事情"不合逻辑"，其实是因为并不理解隐藏在这些事情后面的另套"逻辑"。不但数学家有数理思维逻辑、诗人有诗歌思维逻辑，即使"思维混乱"的精神病患者仍有着自己的"逻辑"。

中国古代存在传统的发现和论证模式，但没有西方意义上的logos，为了避免含混不清，现在给出一个"发现和论证模式"的上位概念"思辨模式"。语言模式和思维模式的差异性导致了思辨模式的不同性：屈折结构类型产生了以同一律为原则的严密（分析性、缜密性）而求实（实证性）的"逻各斯"，其典型格式是三段论。词根结构类型产生了以同构律为原则的周详（轮廓性、整体性）而达意（意会性）的"意象法"，其基本习惯是推类法。

闪含语和印欧语以语音屈折变化为手段表示语法意义，导致了语音音素的抽象分析，形成了以"分析——归纳——演绎"为过程的"概念——判断——推理"的思维模式。分析法与音素分析、归纳法与音素归类、演绎法与音素组合，分别具有相关性。从概念到判断是归纳过程，从判断到推理是演绎过程。在三段论中，大前提是归纳的结果，小前提是个别的判断，结论是演绎的结果。

希腊与印度虽然相距遥远，但是各自产生的思辨模式都不约而同地趋于三段论。德谟克利特（Demkritos）将思维作为研究对象，写出了发现归纳法的西方《逻辑》。发现音素的思维基础就是自发的归纳法，因此作为源自音素论的原子论的创立者，德谟克利特对归纳法的发现不可能与音素论没有联系。亚里斯多德在《工具论》中提炼出由大前提、小前提和结论三部分组成的演绎式三段论。经逍遥派、麦加拉派、斯多噶派的补充和发展，古希腊的形式逻辑建立起来了。在印度因明学中，思辨模式

提炼为宗(断案)、因(小前提)、喻(大前提)组成的“三支式”。三支式与三段论名异而实同,只是内部排列顺序有别。半个世纪以前,虞愚曾经发出慨叹:“亚里斯多德所著《机关》一书,其中讨论结论与验证之‘分析法’,及讨论归纳及或然之‘辨义法’与印度因明学属有统系之研究,为古今东西学者所共尊,谅有多少关系,不尔,五分作法与《机关》之三段论式,何相似之甚耶?(此问题惜无确实考证犹属悬案)”[①]。依据三支式源于《正理经》的五支式(宗因喻合结),可以认为印度因明是自源逻辑。三支式与三段论思辨模式相似性的“悬案”,只能以语言类型对思维模式的制约性原理来加以解释。

最早的一批阿拉伯逻辑家都是古希腊逻辑著作的译注者,他们的功绩是结合阿拉伯当时的情况宣传了逻辑的作用。除了使古希腊逻辑延续下来,后来又把逻辑论著反传西欧[②]。阿拉伯逻辑学之所以创新不多,是因为希腊逻辑适合于他们的语言类型及思维模式,不可能有质的突破。尽管西亚没有产生本土形式逻辑,但是依据印度自发产生本土因明三支式的事实可以推断:如果西亚自发产生逻辑,也许叫别的名称,但必然是“三段论”。依据语言类型对思维模式的制约性原则,可以给出如下三段论:“如屈折型语言的思辨模式皆表现为三段论;如闪含语言为屈折型语言;闪含语言的思维模式是三段论”。

只有词根型的华夏汉语与之不同,一方面语义与语音合一为“名”,另一方面通过“名”的直接配比来传情达意,导致了语义的具象训释,形成了华夏汉人以“立象——同构——顿悟”为过

① 虞愚《因明学》,中华书局1989年影印版。《机关》(Organon)今译《工具论》。因明的“五分作法”即五支式,后改为三支式。

② 宋文坚《西方形式逻辑史》,中国社会科学出版社1991年。

程的"意象——比类——尽意"的思维模式。以华夏汉语为参照，立象与具象命名、类比与同源派生、顿悟与转注立训，各自具有相关性。以华夏汉字为参照，立象与象形造字、类比与假借用字、顿悟与形声造字，分别具有相关性。从意象到比类的过程是同构，从比类到尽意的过程是顿悟。这种思辨模式可称之为意象法，其典型格式可称之为推类法。

中国人的认知思维中，最基本的概念是"象"。《周易·系辞》曰："天垂象，……圣人象之"。"垂象"之"象"是名词，"象之"之"象"是动词，圣人象象。又曰："仰则观象于天，俯则观法于地，观鸟兽之文与地之宜，近取诸身，远取诸物，于是始作八卦，以通神明之德，以类万物之情"，"象、法、文、宜(仪)"皆为象之对象。华夏汉人以"意象"来表达事物的性质，与"类比"和"尽意"组成思维模式。意象(象象)是对"象"的规定，即如何依据个别事物的显性特征将客观存在转化为主观认识，实现的办法是"观物立象"或"取象"。比类(类象)是对"象"的系联，即通过具有相似性或相关性的不同事物的系联将个别认识扩展为一般规律，实现的办法是"援物比类"。尽意(言象)是对"象"的阐释，即在"观物立象"和"援物比类"的基础上，选择合适的修辞表达顿悟的感受[①]。"圣人立象以尽意"，在中国古代文化中，"取象"法普遍存在。汉字起于象形，故《说文》取象于"象"以统文字。八卦取象于男女交合，故喻之"乾坤"连断。太极取象于男女，故图以"双鱼"互补。《周易》取象于"易"(蜥蜴)以比变易，老子取象于"道"(产道)以示有无，孔子取象于"仁"(二人)以明忠恕，庄子取象于"鲲鹏"以求畅达。

---

① 参见徐通锵《语言论》，东北师范大学出版社 1997 年。

胡适认为《墨子》的"三表法"是先秦名学的论证理论[①]。《墨子》说："凡出言谈，则必先立仪而言，是故言有三法(与"三表"同)。何为三法？曰：有考之者，有原之者，有用之者。恶乎考之？考先圣大王之事；恶乎原之？察众之耳目之情；恶乎用之？发而为政乎国家，察万民而观之"。如同有人要把《周髀算经》译解成欧氏的《几何原本》一样，也有人试图把《墨经》译解为亚氏的《逻辑学》，其实《墨子》的"名、辞、说"与亚氏《逻辑学》的"概念、判断、推理"的组织形式并不一样[②]。三表法是三种单独的也可以同时使用的思辨方法：或法圣王(给出权威论据)、或察耳目(印证亲身体验)、或观为政(表明主观动机)，它们之间并不具备三段论内部的推理联系。究其"必先立仪"[③]，其思维模式还是比类以尽意。

朱晓农认为，与西方形式逻辑所遵循的同一律不同，秦人思维的基本观念是同构律，其推理方法是与具有盖然性的归纳法、具有必然性的演绎法和具有偶然性的类比法都不同的内推法、外推法和旁推法。同构律是建立在对种—属、属—种和异类三种同构性的认可之上。其中异类同构性及其旁推法是秦人同构律的认识论和方法论的核心[④]。"万物同构观"的出发点就是万物的相似性，也就是在不同的事物之间寻找共同性或相关性的以此推彼，因此最典型的方法就是推类法。推类法没有大、小前提和结论的三段式，只要列出异类的同构性，就可以断言甲类具

① 胡适《先秦名学史》(李建钊等译)，66页，学林出版社1983年。

② 汪奠基《中国逻辑思想史》，上海人民出版社1979年。

③ "仪"、"宜"同词，即"义"、"意"，与"象"同义。

④ 朱晓农《秦人逻辑的任意性和旁推法的两种推理模式》，《走向新世纪的语言学》，台湾万卷楼图书有限公司1997年。朱另有《为什么中国古代哲学家没有发现三段论?》，《自然辩证法通讯》1986年第1期。

有乙类的特性。

根据西方的三段论"如果所有的人都是有死的,如果张三是人,则张三是有死的"为标准,中国式的推理"人固有一死,圣人人也,然则圣人不朽"是不合"逻辑"的。其实西方三段论的对象是"生物性的人",死是"生物性的死";而中国式的推类所讲的是先是"生物性的人"和"生物性的死",后转到"文化性的人"和"文化性的死"。这种不保持共同性、强调独特性的旁推法就是古代中国思想家、语言学家、历史学家常用的推类方式之一。[①] 作为生物性的人必然会死,作为文化性的人可能不朽。正如一首诗说:"有的人活,已经死了;有的人死了,他还活着",就是着眼人的精神文化价值的不同。西方的三段论形式上很科学,保持了概念的同一性,但是却没有揭示人的文化性,在中国古代学者来看三段论所论皆为常识,句句是"废话"。中国的旁推法形式上违背了同一律,"偷换"了概念,在西方形式逻辑来看无异于"诡辩",但却提供了新知。根本问题在于,形式逻辑的同一律掩盖了事物属性的复杂性,而逻辑只是人们认知世界的一种证明手段。"概念——判断——推理"的思维模式和思辨模式三段论,以事物的同一律为原则,严密分析论证,重在"以理论之";"意象——比类——尽意"的思维模式和思辨模式推类法,以事物的同构律为原则,整体感受顿悟,重在"以心悟之"。然而不但异类事物是同构性和非同构性的统一,即使同类事物也是同一性和非同一性的统一。西方概念重在具体属性的单一性和不同属性的相异性,中国意象重在异类对象的整体性和不同事物的相似性。属性的单一性与音素的性质相同,对象的整体性与音节的性质相同。不同的语言类型与不同的思维模式具有相关性,这就是

① 朱晓农《秦人逻辑的任意性和旁推法的两种推理模式》。

东西方思维具有不同特点的内在根源。

## 七、希腊原子论模式与音素论模式的同构性

本原哲学的对象源于原始神话传说母题。生态环境造成了蓝色文明与黄色文明的不同，从认识对象上造成了西欧源于自然神崇拜的“自然哲学”与华夏源于祖先神崇拜的“人伦哲学”的差别。本原哲学的模式源于语音结构模式。人们通常以为“希腊哲学是西欧学科之母”，而不知哲学是本原语言文字（书语）和原始神话传说（口语）的智慧女儿，远远早于本原哲学的语言文字学才是人类学科之母。

距今大约 6000－7000 年前，从原始记事符号中孕育出最早的文字。文字的发明为巫术文化、社会管理、历史记载和商业会计提供了超时空的传播工具，使人类进入了依据文字“垂后识古”的文明阶段。人类文化是对周围环境的适应，新石器时代的每种语言都受控于特定的生态环境，每种语言都在以特定的方式切分周围的世界。运用的原初语言不同，对世界的原初认识也就有别，由此形成特定的原始认知模式。自源文字的发明和使用必然以原初语言的天然感知和直觉分析为前提，因此对语言进行分析的自源文字成为人类文化的原初学科。从宏观文化的视界来看，与其他文化不同，只有语言才具有本体文化和载体文化的双重性。用语言文字做载体的后起学科势必留下最初的语言感知和直觉分析的烙印，从而导致了哲学思维模式与语言结构模式的同构性。

希腊本原哲学既是脱胎于自然神崇拜的“自然哲学”，又是以屈折语的音素分析模式为基础的“物理哲学”或“分析哲学”。依据 word 内部音素结合继起性的离散式知觉，古埃及含米特人轻而易举就发明了分析辅音的“离素原则”，制定了最早的辅音

符号。与之语言结构相同的闪米特人，通过借用音素符号制订了辅音字母表。依据印欧语的屈折特点，希腊人又进一步完善为辅一元音字母表，形成了把音系分析归纳为有限音素、由这些音素再组合成若干音节的“分析归纳——组合演绎”的思维模式。与之相应，“元素—原子论”把自然界分析为基本元素，用这些元素的结合与分离解释万物的生灭现象。也许希腊“元素说”自然哲学家泰勒斯(Thales)正是从埃及人那里接受了水是太初存在的观念而提出“水一元论”。同样埃及人对太阳神、大气神和土神的崇拜，启迪恩培多克勒(Empedokles)提出了“火、水、气、土”是万物之根的“四元素说”，同时“爱”(吸引)和“斗”(排斥)这两种力量使元素结合或分离。希腊“原子论”哲学的创始人留基波(Leukippos)的传人德谟克利特，进一步提出世界的本原是“原子”和“虚空”。原子这一基本哲学概念，其本义即“不可分的物体”。原子之间只剩下形状的不同，以及数量、位置、次序和彼此的结合与分离。原子在虚空中彼此结合而形成涡旋运动，从中产生出万物。大千世界的生生灭灭、千变万化，都是原子的合与分运动造成的。当代学者仍然认为：古典原子论在许多方面是柏拉图之前的希腊哲学成就之冠。即使在柏拉图和亚里斯多德以后，通过伊壁鸠鲁和卢克莱修，在希腊思想中仍然起着十分重要的作用。

希腊古典时期末期的普利西安(Priscian)在论述拉丁语史所说：“正如原子集合起来产生一个个有形的物体一样，语音组成了像某些有形实体一样的语言”[1]，这一论述已经揭示出原子论与音素论之间的相似性，但未能正本清源。字母文字是在前

[1] 转引自[英]罗宾斯(R.H. Robins)《简明语言学史》(许德宝等译)，30页，中国社会科学出版社1998年

9世纪从腓尼基传入希腊的，原子论哲学则产生于前5世纪左右。依据先后顺序，是音素论模式促成了原子论哲学的产生，而不是相反。语言和文字是希腊古典时期人文教育的基本内容。斯多噶学派认为：首先产生印象，然后是思想借助语言，以语词的形式表达由印象产生的经验；其次，一切事物都是通过辩证的研究来区分的；再次，大多数人同意，辩证法的研究应该从与言语有关的部分开始。哲学的原子论使古希腊确立了以自然客体为认知对象，依据基本元素或原子的离析和结合方式对客体进行分析—归纳—演绎的科学主义认知传统。

阿拉伯人来到叙利亚和美索不达米亚之前，这一地区的许多城市已经希腊化。阿拉伯哲学除了从波斯和印度哲学中吸收了一些思想，主流是对希腊哲学的继承和发展。一方面受原子论和毕达哥拉斯数论的影响，产生了中世纪阿拉伯自然哲学学派；另一方面受亚里斯多德学说影响，产生了阿拉伯理性哲学学派。哲学名词"焦赫尔"(jauhar，本质)被认为是希腊语"原子"的译语。8世纪的巴士拉学派不仅继承了希腊原子论的传统，而且提出了展延性与实体化、位置和地点、物理的可分性与数学上的可分性、数学上的原子与物理学上的原子等概念的差别。从拉齐(Razes，864—925)开始，原子论在穆斯林神学家之间极为流行。拉齐认为，物质是由不可分的原子构成的，这些被太虚所间隔的原子的疏密度决定了客体的基本的质，与德谟克利特学说的相似性非常明显。几乎作为独一无二的例外，穆斯林神学家接受了物质、空间和时间的原子论观点，在它上面建起了真主作为绝对主宰的神学大厦。阿拉伯哲学融合了希腊哲学，吸收了印度与波斯哲学，调和了哲学和宗教，沟通了哲学和科学。反过来，中世纪的阿拉伯哲学又对欧洲哲学产生了巨大的影响。它把希腊哲学保存下来，使处于黑暗时代的西欧学者在读到由

阿拉伯文重译为拉丁文的著作后才发现了辉煌的希腊哲学时代[1]。

吠陀时代的印度就开始专门研究语音规则,早在希腊人之前就创造了语音分类原则,区分出元音与辅音、塞音与擦音。据说公元前1000年就编成了第一批词典,分解出词根、后缀和词尾。印度人把语法称为"语法分析",其意为"分析、切分"。前4世纪的巴尼尼《语法》,包含3996条规则。这些研究使印度的宗教与哲学一开始就带有语言分析的特点。佛陀时代六大师之一的阿夷陀(Ajita Kesakambali)主张"土、水、风、火四大说"。在此基础上,波浮陀迦旃(Pakudha Kachchayana)提出"七要素说","土、水、火、气"是物的要素,"乐、苦、灵"是心的要素。七要素是稳定不变、无生无灭的,一切众生有命的生死都是要素的集散离合现象。耆那教派提出的与原子论相近的"极微说",在列国哲学六宗时代颇为盛行。弥曼差派认为,世界不依赖任何神力而永恒存在,万物的生灭是由于极微变化的结果。数论派的因中有果论,主张"因"是物质的,称之为"自性"或"原质",原始物质细微不可见,故又称之为"未显"。胜论派与正理派的哲学观点都是极微说。据说胜论派的自然主义实在论即源于七要素说,他们认为"微尘是常,能生一切物"。与德谟克利特把灵魂也看作是由原子构成的古希腊一元原子论有所不同,印度的极微说认为"我"(灵魂)是一种独立于物质实体之外的实体,表现为二元原子论[2]。佛教教派的吠婆师迦和经量部,婆罗门教派的正理派和胜论派,还有耆那教,到5世纪都推断出原子论,显然与

---

① 参见李振中、王家瑛《阿拉伯哲学史》,61页,北京语言学院出版社1995年

② 参见崔连仲《从佛陀到阿育王》,辽宁大学出版社1991年。

古希腊哲学的影响无关[①]。为什么印欧哲学中不约而同地产生原子论，这只能用音素论来解释。

西方自然哲学的原子论与语音分析的音素论的同构性在于：1. 因子性，事物由不可再分性的基本单位构成；2. 离析性，这些基本单位之间可以相互离析；3. 组合性，这些基本单位之间可以相互组合构成事物。总之，在西方文化中，世界的统一性在于事物的因子性，事物的多样性在于因子的组合性。

## 八、中国玄牝象模式和音节论模式的同构性

中国本原哲学既是源于祖先神崇拜的人伦哲学，又是以词根语的音义合一性感知为基础的顿悟哲学。华夏汉语是以没有形态变化的"名"（后世统称"字"）为基本符号的词根语，音节内部的音素结合方式具有混成共起性。华夏汉语的派生词以同源孳生为主，孳生词与本源词的关系或是语音全同，或是语音屈折，或是辅音附加，由一个单音词孳生出来的依然是单音词。与古埃及发明的离素原则不同，华夏汉人发明的是以同源派生为基础的假借原则（早期同源词共用字根）和为区别同形字而衍生的形声原则，这两个原则的基础都是语音认同的联想性类比。

中国文字的产生和发展与华夏汉人对语言的感受相一致。《说文》："文，错画也"，交错画文也。"文"即纹身、纹饰之本字。古人从以色彩线条纹人之身、纹器之饰，发展到以"文"记事。《说文》："字，乳也"。段注："人及鸟生子曰乳"。本义是在屋里哺乳孩子，与表滋长、滋乳之义的"兹、滋、孳"等同源。"文"为

① 参见［美］马吉德·法赫里（Majid Fakhry）《伊斯兰哲学史》（陈中耀译），47页，商务印书馆1996年。

"字"之本，则"字"生于"文"。"文""字"孳生如同"母""子"化育，一生就是一个音义兼备的整体。与之相应，先秦老子哲学的创世模式是："有物混成，先天地生"，犹如《说文序》："文者，物象之本"；"道生一，一生二，二生三，三生万物"，即《说文序》："字者，言孳乳而浸多也"。

（一）释"道"

作为"道德哲学"的老子哲学，"道"通常被后人误解为"道路"，引申为"途径"与"方法"[①]。道之"路"义，并非其原始义。考"道"之字形，右"首"为头，左从"辵"为足。《说文》："辵，乍行乍止也，从彳止"；又"彳，小步也，象人胫三属相连也"。依据甲金文，"止"为脚形，"彳"为"止"之变体。"止"与"彳"为双足前后移步之形。"道"有上首与下肢，像人之形。从语源考据，"道"之初文"首"，"首"即头也，为人之主体；"首、道、头"古同音。依据《老子》"道"指婴儿之形。周代金文"道"，从"首"在"行"中，为"衜"。《说文》："行，人之步趋也；从彳亍"；又"亍，步止也"。"彳"与"亍"皆为"止"之变体，"行"为双腿之行。又《说文》："胻，胫端也"，"行"即"胻"之初文。"行"为双腿之形，又通行走之径。甲骨文另一字形 "卝" 盖取象于行走之交叉路口。此为同词而异象也，许慎说解则据其本义。

考定"道"之本义为"婴儿"，除了依据《老子》文本，还依据《说文》"道"之异体作"𨖍"。《说文》："寸，从又一。又，手也"，"寸，度之，亦手也"，"寸"为"肘"之初文。"道"之异体"𨖍"，依然表示"道"之本义与人形有关，而与道路无涉。无论右"首"左"辵"，还是左"首"右"寸"，"道"皆为有首、有肢之人。婴儿顺产时"首"必先出，肢体"肘"和"辵"而后出，故"道"象赤子之形。由

① 任继愈《中国哲学史(一)》，44页，人民出版社1979年。

此可见："道"作为名词指"产道"（老子"玄牝"），作为动词指"产子"。《说文》："導（导），引也"。从道、从寸，以手引子"接生"谓之"导"。"道"之路义为产道、通道的后起引申义。

（二）释"道、达、泰、大"

又"道"有通达之义，"達"之形义也为生子。《说文》："達，或从大"。段注："此与水部滑、泰字音义皆同"。《说文》："滑，利也"；又"泰，滑也。从廾水，大声"，段注："从廾水，水在手中，下溜其利也"。"达"之字形，"大"为人形，"辶"为足形。"泰"（𣴎）为"大"顺水（水）滑下，以双手（𠬞）接之之形也。"大（大）"（人展肢之形）、"达"（顺产之形）、"泰"（接生之形）与"胎"（腹中胎儿）同源。《诗经·生民》叙述后稷诞生时"先生如达"。"先生"，头生子也；"达"，顺产也。此可证"道、达、泰、大"皆为生子之形义。

（三）释"德"

《说文》："德，升也"。段注："升当作登。迁，登也……登读言得……今俗谓用力徙前曰德"。"德"之初文为"悳"。《说文》："悳，外得于人，内得于己也；从直、心"。甲骨文从目从心会意，以目视之而心有所得也。"德"、"得"音同义通。《说文》："得，行有所得也；从彳，㝵声"。初文为"㝵"，甲骨文从贝从寸会意，以手持贝而有所得也。老子之"德"，即母生子有所得也，俗语"喜得贵子"。

（四）释"有、无"

老子又曰："天下万物生于有，有生于无"。"有（𠂇）"，从又（又）、从肉（肉），源于"又"。《说文》："肖，骨肉相似也；从肉，小声；不似其先，故曰不肖也"。"肉"，父母所生之骨肉也。"有"之形以手持肉，即抱子也，生子为"有"也，俗语怀孕叫"有了"。"无"，古音近"母（母）"，"母"之同源有"毋（毋）"，与"无"音同义

近而同源。“无”亦作“無”,与“舞”同源。“舞”与“巫([illegible])”同源,女巫以舞迎神也。“无”、“毋”、“母”、“舞”、“巫”皆同源。所谓“有生于无”,即取象于“子生于母”。“有”包含主观拥有、客观存在二义。“有”与“无”相对,拥有之语义。“有”与“亡”相对,存在之语义。《说文》:“亡,逃也;;从人乚(隐)”。“亡”与“无”为阴阳对转。中国人的客观存在的语义源于主观拥有的观念。主观拥有的另一对概念是“得”与“失”。《说文》“失,纵也。从手,乙声”,从手中失去。

解释了核心概念“道”、“德”、“有”、“无”,老子要义才能畅达。“谷神不死,是谓玄牝;玄牝之门,是谓天地根”。“谷神”与“玄牝”呼应,“谷神”即欲神,即孔子的话“食色,性也”,“玄牝”即女阴,以之喻为“天地之根,众妙之门”。又曰:“有物混成,先天地生。寂兮寥兮!独立而不改,周行而不殆,可以为天下母。吾不知其名,字之曰道,强名之曰大”。先天地生者为天下之“母(毋—无)”,称之为“道”(产道),强名之为“大”(舒展四肢之人)。“有物混成”为阴阳交媾,“独立不改”、“周行不殆”为周期月经。以往的哲学家因为不了解“道、德、有、无”之语源本义,或误解“道德”,或认为《老子》的“有无”是“最概括,最抽象,最难懂的一部分”①,以致于对老子哲学的阐释总存在着一层隔膜。

与希腊原子论脱胎于自然崇拜不同,道德哲学的原型是华夏先民的女性生殖崇拜或祖先崇拜。老子取象于母之“玄牝”化育,援物比类天地万物之“有生于无(母)”,以尽“效法自然(自生自灭之状)”、“无为(无须人为干预)而无不为(无一不自为)”的“君王南面之术(宗法社会的组织管理)”之意。由个体化育到万物造化,再到人伦关系,建立了道德哲学、万物哲学(自然哲学)

① 冯友兰《中国哲学史新编(二)》,48页,人民出版社1984年。

与人伦哲学(权术哲学)的三位一体,展示了中国古代哲学"天人合一"的人伦主义性质。人的产生是阴阳交媾、完形整魄,老子的哲学原则是混成性(阴阳交媾)与整体性(灵肉合一),因此,老子哲学又可以称之为玄牝象,由此促成了取象类比、言诰尽意的顿悟性哲学传统的确立。庄子要讨论精神自由,并不取概念、给定义、下判断与做推理,只是取譬鲲鹏"水击三千里,抟扶摇而上者九万里"而尽"无拘无束"之意。老子玄牝象道德哲学的基本原理与华夏语言文字的孳乳原理具有对应关系:

| | | |
|---|---|---|
| 道德哲学原理 | 母(玄牝)——产生——子(道)<br>无(天下母)—产生——有(德)<br>无为——　产生——无不为 | 取象:生殖化育<br>比类:万物生成<br>尽意:南面之术 |
| 语言孳乳原理 | 原生词——比类孳乳——次生词 | 取象——同源——转类 |
| 文字孳乳原理 | 独体文——比类孳乳——孳乳字 | 象形——假借——形声 |

中国的老子玄牝象结构模式与语音分析的音节论模式的同构性在于:1.意象性,通过意象认识事物的属性;2.整体性,单个事物是整体的、混成的;3.孳生性,万物生于一,事物之间相互孳生;4.类比性,异类事物之间普遍联系,可以连类而及。总之,世界的统一性在于事物的同构性,事物的多样性在于事物的孳生性。

《周易》为西周占卜之书,其哲理是"易"(昜,蜥"蜴"之本字),取象于蜥蜴之易色,以比类人事之变易,而尽万物变易之意。"周"一语双关:一为"周国",一为"周遍"。万物皆变,周而复始,方可占卜。《易》之八卦,基本符号是乾一、坤－－。其"乾坤观"取象于男女、乾坤交合孳生,衍为六十四卦。中国古代虽有貌似二分的"阴阳观",最典型的就是太极图(原始生殖象征的

黑白双鱼图案)，但意在互补而并非离析。乾坤阴阳观来自男女交合和日月交替，日月交替为昼夜，男女交合生子女，仅此二分仍合二而为一，出发点重在事物的互补完整性而不是对立分析性。这种“合二而一”(“中庸”、“尚和”的思想导源于此)的互补性思想模式，仍然与汉语语音和语义“合二而一”的音节论相一致，与西方音素论的音素多元性与组合多样性的“一分为多”思维模式具有本质的区别。乾坤阴阳观的特征仍然是意象性、整体性(互补性)、孳生性和类比性。

季羡林先生是深知“语言文字是思想的外在表现形式，而思想的基础或出发点则是思维模式。东西方的思维模式是根本不同的。西方的思维模式是分析。……而东方的思维模式则是综合，其特色是有整体概念和普遍联系的概念”[1]，遗憾的是，在东西方文化的划分上却没有贯彻这一准则。

## 九、西方演绎科学文化与中国经验技术文化

语音模式、哲学模式、思维模式对子文化结构模式具有制约性，子文化结构与语音模式、哲学模式、思维模式具有相关性或深层通约性。同一类型的各个子文化模式是语音分析模式、本原哲学模式与认知思维模式在这些衍生文化领域内的历史性自然延伸。古希腊 logos 的最初含义是“讲话”(话语表达)，后来才专指“理性”(理性分析)。“神话”(muthos)一词开始并不与 logos 对立，其原义也是“叙述”(口头传说)。前 6 世纪，泰勒斯等人创立了把自然当作对象，探索宇宙的起源和构造的一种新的思维方式。随着古希腊元素—原子论自然哲学的诞生，逻各斯仿佛突然脱离了神话，出现了人类思想史上的“希腊奇迹”。这

1. 季羡林《中国现代语言学丛书·序》，4 页，东北师范大学出版社 1996 年。

些最初的自然哲学尚未完全蜕尽神话结构的胎记，不仅宇宙演化论重复并延伸着神话中的宇宙起源母题，而且各种自然"元素"在失去人格化的神的特征的同时，运作中仍然充满了那种曾为宙斯所特有的智慧和公正。

前3世纪，希腊自然科学终于实现了全新的飞跃，走上了一条演绎科学的道路。这门随着欧几里德（前330～275）理论而出现的、涉及抽象对象的几何学，从有限的几条公设、公理和定义出发，把严密演绎得出的命题一个接一个连接起来，用具有形式化特点的验算来保证命题的有效性，并在连贯的推理中把这种有效性确定下来。与只要说明程序的古埃及土地测量方法的发现不同，几何学成为形式化证明的抽象科学。古希腊几何学不仅是亚里斯多德建立逻辑学的学术前提之一，甚至决定了亚里斯多德的科学观[1]。以此为基础，希腊自然科学形成了"原理（直觉归纳）——假设（演绎推理）——证明（实验演算）"的研究特点，形成了延续至今的欧洲科学主义传统。虽然匈牙利的萨博（A.Szabo）和英国的劳埃德（G.E.R.Lloyd）等科学史家认为希腊人在数学上的发展完全处于希腊思想变革的范围之内，但是赛敦博格（A.Seidenberg）从前5～前2世纪的印度教科书中，发现了建筑祭坛时采用标准的规定。他认为，毕达哥拉斯定理在这些规定中所占据的位置表明，巴比伦的代数学、希腊和印度的几何学有一个共同的来源[2]。公元7世纪以后的阿拉伯人主要通过波斯学者的传介接受了印度的医学、数学、天文学。印度的天文学和数学又启发阿拉伯人翻译了欧洲托勒密的天文学著作《大综合论》与欧几里德的《几何学原理》。阿拉伯人还接受了印度数字，再传入欧洲被称为阿拉伯数字，为西方数学的符号变

---

[1][2]　[法]皮埃尔·韦尔南《希腊思想的起源》，12－14页。

革提供了条件。尽管"希腊奇迹"如此神奇,欧洲自然科学的成长还是萌芽于古代埃及的学术。从自然科学来说,埃及人首先创立了数学和机械学,发明了化学和医学;从语言学来说,埃及人首先发明了语音符号和语音分析的"音素论"原则,为"元素—原子论"的创立提供了理论模式。

虽然范德瓦敦(B.L.Van der Waerden)认为,依据中国古代的算术论著与巴比伦的算题汇编的相似性,表明它们可能来自同一源泉。但是,"算术"(具象的"算"与抽象的"数"不同,技艺的"术"与研究的"学"有别)在中国始终没有成为一门显学,更没有以数学思维为原则贯通其他知识门类形成西方意义上的自然科学。古代中国的科技主要是经验性的记载、应用性的技艺发明,泛称为"术",最富特色的就是"四大发明"。就主要倾向而言,中国的"化学"是为求长生不老的道家"炼丹术";"天文学"是星相的观察与历书的编纂;"地理学"是疆域的划分、风水的"方术"、军事的"方舆"和文学的"游记";"生物学"是植物的栽培和动物的豢养;"生理学"是道家的"房中术"和医家的"筋络说";"医学"是草药和"针灸术"。西方的几何学、物理学等等,是在明末由西洋传教士利玛窦等人带来的,其他一些自然科学是在西学东渐时的舶来品。在中国,无论是人文学术还是应用技艺,占主导地位的都是以音节论和意象法为模式的"联想—类比"或"整体—顿悟"的人文主义传统。

## 十、西方复式焦点艺术文化和中国单式散点艺术文化

艺术文化至少应当包括文学艺术、音乐艺术和美术艺术。文学艺术传统模式的比较拟另文讨论,这里主要就音乐艺术,参对美术艺术,探讨东西方传统艺术模式与东西方语言模式、思维

模式和哲学模式的相关性[1]。

（一）中国乐音的相对性观念与西方乐音的绝对性观念，与东西方不同的语音结构具有相关性。

由于汉语是单音节声调语言，声调具有区别字义的功能，乐音结构的变化与调值的平仄升降务必力求一致性，由此形成了乐音的相对性观念。在“音腔”特色的基础上，衍生了腔格、腔韵、腔句、腔调、腔套、腔系等音乐结构层次。由于西方语言是非声调语言，音节的高低升降不会引起词义的变化，乐音结构的变化无须考虑语词音节高低升降的变化，由此形成了乐音的绝对性观念[2]。与中国乐音结构的波动式特征不同，西方乐音结构的特色如同西方语言句调的升降，具有明显的阶梯式特征。

（二）东西方音阶旋法和音乐律制的不同，与东西方思维模式的整体性和分析性的本原差别具有相关性。

就旋法而言，东西方音乐表现为中国“无半音五声性旋法”与西方“有半音七声性旋法”的对立。中国传统为五声音阶，五声为“正声”，五声之外称为“变声”。印度、西亚、欧洲的传统是七声音阶。乐音的全音和语音的音节具有相似性，与中国人对于汉语音节的自然感知和整体把握相一致。乐音的半音和语音的音素具有相似性，与西方人对于音素的自然感知和离析切分相一致。全音音阶的思维模式是音节论，半音音阶的思维模式是音素论。半音的产生意味着全音内部的再分析，为不同调式的生成提供了音阶条件。旋法是律制的显示和使用，律制是旋法的提炼和原则。就律制而言，东西方音乐表现为中国的三分损益法与西方的十二平均律的对立。中国先秦时代的律制是

---

[1] 音乐文化的内容，听取了南京师范大学音乐系民乐家岳峰教授的意见。

[2] 参见王耀华《世界民族音乐概论》，上海音乐出版社1998年。

"三分损益法"(《管子·地员篇》),由此产生五声宫调。古希腊的毕达哥拉斯以数学方法,使用弦长进行精密推算,在十二平均律的基础上产生和声。虽然公元前2世纪,中国在音乐实践中出现了偶尔十二平均律的使用,但是直至明代朱载堉的《律学新说》(1584)才提出平均律理论,其来由可能与佛教音乐或西洋乐理的传入有关。中国的律制为经验测定,与中国的技艺发明相一致;希腊的律制为数学推算,与西方的科学方法相一致。

(三)东西方音乐节律的不同,与东西方语言结构和经济方式的不同具有相关性

不同的音乐文化都有自由节律(非均分律动)和均分律动,但中国以自由律动为主,西方以均分律动为主。即使同样使用均分律动,东西方仍然存在"松紧式均分律动"和"强弱式均分律动的差别"。中国是以二拍子为基础的均分律动体系,这种律动的相互交替,与其用"前强、后弱",不如用"前松、后紧"来描写更为贴切。欧洲和西亚具有代表性的是功能性的"强弱式均分律动",波尔卡的每小节是强、弱,圆舞曲的每小节是强、弱、弱。有人认为,二拍子可以说是农耕文化的劳动号子的呼应节奏,以稳定的平面活动为基础的"一松一紧"的节律;三拍子可以说是游牧文化的骑马奔腾式的起伏节奏,以颠簸的上下跳动为基础的"强、弱、弱"节律[1]。显而易见,劳作节奏对音乐律动的形成具有一定的影响,但是还必须考虑到语言节律对音乐律动的差异所产生的必然作用。从发生学来说,语音和乐音具有同源共生性和结构相似性,乐音只是特殊的或专门化了的"艺术语音"。音乐节律以语言节律为基础,正像语音中的规整节律是从自由节律中发展起来的一样,乐音中的均分律动也是从自由律动中发展起来的,相互影响自然

[1] 参见王耀华《世界民族音乐概论》

而长久。松紧是气流(重在速度)的舒缓控制,强弱是气流(重在力度)的强弱的控制。上古汉语的音节结构是元音后跟塞音尾的闭音节,音节内部元音舒缓而塞尾短促。中古诗歌的平仄律,平声舒缓而仄声顿促,有人称之为长短律。与前松后紧的汉语语音特点不同,印欧语言单词具有词重音,重读音节强而弱读音节轻。均分律动的前松后紧差别有如汉语音节的前舒后促(上古汉语音节结构为闭音节),均分律动的前强后弱差别有如西方词音的重读弱读。乐音律动和语音律动具有明显的相关性。

(四)不同的曲式和织体受控于不同的乐思,不同的乐思则是东西方思维模式在音乐领域的自然延伸。

乐曲包括曲式、织体和乐思三个层次。就曲式而言,东西方音乐主要表现为中国的"渐变衍展式"与西方的"顿变对比式"风格的不同。中国音乐曲式以 A B C(或 $A_1$ $A_2$ $A_3$)的持续展开为典型,不再返回原处,形成渐变衍展的风格。欧洲音乐以三部曲 ABA 最为典型,以 B 的顿变与 A 对比,和 A 的再现来呼应,形成顿变对比的风格。就织体而言,东西方音乐表现为中国的横向单织体与西方的纵横复织体的不同。中国传统织体是以横向结构为基础的单声部的单织体。西欧乐曲是注意横向旋律和纵向结构、伴奏多采取和声进行的多声部的复织体。西亚织体虽然以横向性为基础,但由于运用均分律动,旋律节奏和伴奏节奏之间相互对比映衬,又表现为独具一格的纵向性复织体特征[1]。蕴含在曲式和织体深层的是乐思。就乐思而言,表现为中国的散点贯穿与西方的焦点再现的对立。中国的传统音乐以一段曲调为基础,用加花、添眼、换头、衍展等手段,形成一种横向的流变贯穿,如同人在画中游,人移景变,是"散点艺术思维方式"。

① 参见王耀华《世界民族音乐概论》

欧洲的传统音乐用顿变的方式追求鲜明、强烈的变化，以再现的手法构成回旋的主题，如同景为人而设，反复观照，为"焦点艺术思维方式"。阿拉伯风格乐思的特点是单一性和反复性的结合。单一性就是只用一种原型无限地扩大；反复性就是以一种原型的反复形成无限的变化，最终回归到旋律原型。中国音乐从旋律的渐变衍展和织体的单一中追求乐曲的整体连续性，欧洲音乐从旋律的顿变对比和织体的复合中追求乐曲的分析协和体。而这种以有限的几个音编织出变化无穷的旋律图案的阿拉伯乐思结构，与语言的音素论和哲学的原子论模式完全相通。与之相应，东西方记谱方法也各有相同。古代中国创造了律吕谱、宫商谱、减字谱和工尺谱等。西方原始记谱法不详，西欧后来使用的是五线谱和数码谱。中国记谱法是单式结构，和中国传统乐曲的单织体相适应；西方五线谱是复式结构，与西方乐曲的复织体相适应。

（五）传统乐器的起源固然与生态环境和经济方式有关，但是传统乐器的发展变化与音乐思维模式密切相关。

据考古研究，弦乐起源于两河领域和古埃及的原型为弓的竖琴。据文献记载，先秦时期依据制作材料，将器乐分为八音。最近报导，中国境内考古发现的史前土埙（半坡）和骨笛（贾湖），也许是世界考古发现的最早原始管乐之一。如果一旦证实管乐确实起源于东亚大陆，那么东西方器乐的最初差别可能就是原始管乐和弦乐的差别，或许就是农耕文明（利用粘土和管状物）和游牧文明（利用狩猎的弓弦）器乐的最显著的原始差别。就弦乐而言，中国传统为指弹弦乐（筑、筝、瑟、琴），而西方在古埃及时期就发展出弓擦弦乐（竖琴）。就管乐而言，中国最常见的传统器乐是无簧管乐（埙、笛、箫），而西方发展出丰富多样的有簧管乐。与乐曲复织体、和声演奏法的思维模式相一致，西方管弦乐器种类繁多，

为18世纪欧洲大型交响乐的出现提供了工具条件。而中国乐器的相对单调性与乐曲织体、齐奏的思维模式相一致。

（六）中国音乐的宗法教化与西方音乐的宗教信仰的社会功能表现为对立。

从音乐文化的功能来看，中国传统音乐是宗庙祭祀、朝廷礼仪等活动的有机组成部分（周代《乐经》列入"五经"），其社会功能是为源于祖先神崇拜的宗法伦理服务，可以称之为"教化音乐"。外来的"佛教音乐"和晚起的"道教音乐"，主要被用来超度亡灵，消融在中国传统宗法文化之中。西方传统音乐是宗教仪式的一部分，其社会功能是为源于自然神崇拜的宗教信仰服务，可以称之为"信仰音乐"。载歌载舞的古希腊酒神节，祭祀的是大地回春之神狄俄尼索斯。古印度音乐所歌颂的是神、永恒和宇宙这类超越人类的"绝对真理"。基督教的教堂音乐是对上帝的歌颂和心灵的慰藉。文艺复兴以后的欧洲古典音乐虽然吸收了民间音乐的营养，但还是源于中世纪的教会音乐。约而言之：除了民间自娱音乐，中国传统音乐是宗法教化音乐文化，西方传统音乐是宗教信仰音乐文化。东西方音乐文化社会功能的差别与东西方本原哲学内容的差别具有一致性。

（七）东西方艺术文化的差别，不但存在于音乐文化之中，也存在于美术文化之中。

同时，在各自的音乐文化和美术文化之间，可以发现艺术思维的共同点。关于东西方美术构思和艺术手法的差异，美术界向有中国画是散点视而西洋画是焦点视、中国画写意而西洋画写实的说法。焦点透视原理是客体镜象的产物，重在立体，可以说是西洋几何学在美术领域的应用，与西方的理性主义的精微观察相一致。散点意合原理是主体心象的产物，重在印象，可以说是中国神韵说在美术领域的再现，与中国悟性主义的意会感

受相一致。美术与音乐的焦点和散点有相通之处。西洋画的写实技法，与屈折语类型的结构精密与语法的逻辑分析一致，又与其派生的理性科学思维模式同构；中国书画的务虚手法，与词根语类型的结构简约与语义的虚灵感受一致，又与其派生的悟性技艺思维模式一致。西方以油画与雕塑的惟妙惟肖将镜象艺术推到极至；中国以水墨画与草书的了了几笔将心象艺术发展到顶。西方的绘画从19世纪的印象主义，经后印象主义发展成20世纪层出不穷的野兽派、超现实主义等。这些"非理性主义"艺术，高扬否定传统的大旗，不再追求客体的逼真，而是采取形体变形、光色变化、时空叠置等手法表现个人感受的印象流，才实现了从"镜象"到"心象"艺术的转变。通过以上东西方音乐与美术艺术文化的对比分析，似乎可以总结：中国人传统艺术思维的特征为单式连锁性思维和心象艺术风格，西方传统艺术思维的特征为复式综合性思维和镜象艺术风格。

## 十一、文化类型划分的标准与东西方文化的分野

综上所述，世界文化类型划分的标准大约可以确定为四类14项：

（一）源头文化标准

1. 文明起源（新石器时期、金属时代考古文化的同源关系）；

2. 文明类型（与生态环境具有相关性）；

3. 原始崇拜（神话传说的内核，集体无意识的基础之一，人性文化和社会组织文化的底蕴）；

4. 精神信仰（人性文化、道德文化和社会组织文化的心理结构）。

（二）语言结构标准

5. 语言关系类型（史前文明的同源或传播关系）；

6. 语言结构类型(与原始认知模式具有相关性);

7. 语音分析模式(对语言结构的自发性分析);

8. 文字源流谱系(文明传播的重大线索)。

(三)本原哲学标准

9. 思维模式(与语言模式具有相关性);

10. 思辨模式(思维模式的格式化);

11. 哲学模式(与原始崇拜内容相承,与语言分析模式同构、认知对象与思维模式的结晶)。

(四)衍生文化标准

12. 组织模式(个人价值观、社会组织法);

13. 科技模式(语言模式、思维模式、哲学模式在生产与科研领域的延伸);

14. 艺术模式(语言模式、思维模式、哲学模式在艺术领域的自然延伸)。

通过以上各项的比较,可以概括出相关文化各自的总体特征。以几种总体特征不同的文化为基础,可以归纳出不同的文化模式或文化类型。由此可见,西欧、印度、伊斯兰文化具有文明的同源性和结构的相似性,与华夏文化迥然不同。如果暂时不考虑美洲古代文化,那么欧亚大陆的"东西文化"的本质差异及其分野可以图示如下:

| 文化类型 | 东方文化(象性文化) | 西方文化(素性文化) | | |
| --- | --- | --- | --- | --- |
| 文化名称 | 华夏汉文化 | 伊斯兰文化 | 印度文化 | 西欧文化 |
| 国家区域 | 中 国 | 伊斯兰世界 | 印 度 | 欧 洲 |
| 1. 文明起源 | 东亚文化圈 | 地中海文化圈 | | |

续表

| | | | | |
|---|---|---|---|---|
| 2. 文明类型 | 农耕民型<br>黄河长江流域 | 牧畜—<br>通商民型<br>两河流域 | 牧畜—<br>农耕民型<br>印度河—<br>恒河流域 | 牧畜—<br>通商民型<br>爱琴海—南欧 |
| 3. 原始崇拜 | 祖先神为主导 | 自然神为主导 | | |
| 4. 精神信仰 | 宗法伦理 | 伊斯兰教 | 婆罗门—<br>佛教 | 基督教 |
| 5. 语言关系 | 东亚—南洋<br>语言圈 | 闪含—印欧语言圈 | | |
| 6. 语言结构 | 孤立结构 | 屈折结构 | | |
| 7. 语音模式 | 音节论<br>（表意文字） | 音素论<br>（表音文字） | | |
| 8. 文字谱系 | 汉字谱系<br>（自源文字） | 闪米特字母谱系<br>（借源文字） | | |
| 9. 思维模式 | 意象—<br>比类—尽意 | 概念—判断—推理 | | |
| 10. 思辨模式 | 类推法 | 三段论 | 三支式 | 三段论 |
| 11. 哲学模式 | 混成说<br>（人伦哲学） | 极微说<br>（自然哲学） | 原素论<br>（自然哲学） | 原子论<br>（自然哲学） |
| 12. 组织模式 | 集体价值优先<br>集权制·律治<br>（疆域统一） | 个人价值优先<br>分权制·法制<br>（小国林立） | | |
| 13. 科技模式 | 应用技术<br>（经验主义） | 自然科学<br>（科学主义） | | |
| 14. 艺术模式 | 单式连锁性<br>散点意合原理<br>（心象艺术） | 复式综合性<br>焦点透视原理<br>（镜象艺术） | | |
| 地理洲别 | 亚洲 | | | 欧洲 |
| 生态种群 | 蒙古利亚种群 | 高加索种群 | | |

由此观之，季羡林先生的东西方文化的划分是不能成立的。依据欧亚大陆这四种文化的本质差别，合适的划分可能是：中国文化为东方文化，而印度文化、伊斯兰文化与西欧文化为西方文化，至于西方文化内部的差异及其他地区的文化特征可另作别论。借用地理位置来称名仅是出于习惯，也可以用别的术语，如词根语文化与屈折语文化、宗法文化与宗教文化、自然哲学文化和人伦哲学文化、科学文化与技艺文化等等。假如地理位置不仅不能够显示不同文化的类型，而且有可能出现误导，建议选择象性文化与素性文化这两个术语。立象—尽意模式的象性文化，其思维具有整体性、模糊性、比况性等特征；离素—推理模式的素性文化，其思维具有分析性、细密性、实证性等特征。

## 十二、音素论与音节论在人类文化史上的地位

人类学家肯尼思·奥克利认为，现代人的出现是由于语言完善到我们今天这样的水平而引起的一系列连锁反应的结果。换而言之，现代语言造就了现代人，现代语言造就了现代人的思维模式和文化模式。

在东西方文化史上，以屈折语为对象、创立于公元前 2000 年左右的语音分析的音素论，和以词根语为对象创立于公元前 1500 年左右的音节论具有核心地位。音素论模式以因子为基础，认为整体由因子构成，世界的统一性在于因子的物质性；音节论模式以整体为权重，认为部分融合于整体，世界的统一性在于联系的相似性。音素论导致了客体的细密分析，形成了西方的自然哲学和科学主义；音节论导致了主体的整体把握，形成了中国的人伦哲学和顿悟主义。不但东西方的本原哲学结构和对事物的研究方法如此，就是传统的个人价值、社会组织、科技文化、艺术文化等等一系列领域，也正是音素论与音节论的认知思

维模式导致了两种文化类型的分野。

中国学者对语音的分析性研究方法,始于公元前后随着佛法西来的梵文胡书原理的传入。源于西亚音素论的梵文胡书的拼音原理,促使汉人产生了与原初证音法不同的音节二分分析法——反切法,这距离殷商甲骨文时代已有千年之遥。魏晋产生了将韵母归类的韵书雏形。齐梁沈约等人受梵呗新声启迪发现汉语四声,由此形成以调统韵的传统韵书格式。唐宋参对藏文字母,守温创制字母,受印度悉昙影响,胡僧研制等韵,形成以韵图为体裁的音位学。运用音素字母分析书写汉语文献的最早是唐代时期的于阗字母和藏文字母,运用音素字母最早全面分析描写汉语音系的是元代的《蒙古字韵》,制订汉语罗马字记写官话音系的是明末来华的西洋传教士利玛窦,但一直没有引起汉儒的充分兴趣和系统模仿,以致古音学研究一直在音类归纳中摸索,还得等待瑞典汉学家高本汉构拟中国古音。乾隆年间,贾存仁才明确提出汉语音节"头、颈、腹、尾"的四分法,近人刘复补充"神",详尽的声韵调分析法才最终完善。直至西学东渐、西洋传教士制订汉语方言字母和19世纪末的切字运动,才使音素分析法逐步传播[1]。在东西方文化的撞击中,中国人最终接受了西方的音素论思维模式和自然科学研究方法的影响。

长期以往,研究东西方文化的学者趋之若鹜,论著汗牛充栋,一再提出西方"理性思维"的分析性、精密性、实证性和逻辑性,中国"悟性思维"的整体性、模糊性、顿悟性和意合性的特点,但是从来没有进一步提出:是什么缘由形成了这两种思维类型?在以往的东西方文化差异讨论中,一直没有找到一个根本性的

---

[1] 李葆嘉《论华夏汉人的原初音节证音法与反切法的产生》,《南京师大学报》1998年第2期

支点。对于西方本原学术的研究，仅停留在古希腊哲学上面，从未试图进一步揭示西方本原哲学结构模式的由来，更不了解人类有史以来的学术的最早缘起，应当到最初的文字类型即对语言结构的自发分析中去寻找。既然语言不仅是认知世界的主要工具，而且语言本身就是认知世界的主要行为和成果，既然语言模式与思维模式具有相关性、与哲学模式具有同构性，既然人类史以文字的出现而界定文明前后，那么只有揭示出音素论与音节论这一根本差异，才能找到文明史以来东西方文化发展方向不同的本原依据。也正是依据这一线索，有史以来的东西方文化一系列差异的形成缘起得以揭示。如果"打破沙锅问到底"：语音结构的继起性与共起性又是从何而来呢？依据目前历史语言学和史前文化的研究成果，也许只能回答：早期生存环境和原始经济方式，特别是新石器—金属时期的状况，决定了东西方各自祖先的"认知·语言·文化"混沌结构体的走向，最初的微小差异导致了后来逐步形成的不同方向。

对于东西方文化的本原或传统差异的比较，只是一个冷静的、系统的分析，并没有任何情感的贬褒色彩，虽重在揭示东西方文化的本原或传统个性特征，并非否认人类文化具有深层共通性。假如有外星智人来到地球，他们会认为这种差异可以忽略不计而统称为"地球人模式"。尽管如此，东西方文化本质差别的研究，其价值仍然在于为了早日走出困扰了中国几百年的现代化或世界化问题的怪圈。除非被迫中断或者主动毁弃而导致文化转型，在人类文化的长河中，任何类型文化的自在发展都"万变难离其宗"（生存环境最初孕育的文化母胎），现在只是过去的延伸，未来只是现在的发展。如果存在所谓文化基因的移植和交融，也是一个极其漫长而吉凶未卜的曲折历史过程。因此，只有发现了中国文化的生命密码DNA，才能够从根本上解

释3500年来的过去中国、准确描写东西方文化冲突交融中的16-20世纪的近现代中国和预见21世纪世界化进程中的未来中国。

# 跋 走出人文主义与科学主义对立的怪圈

这些文字，是从80年代初以来就理论语言学进行探索的足迹。这些足迹，表现为一个接一个的专题研究，语言的起源和类型、文字的起源和类型、语言符号的可论证性、东亚语言的分布和关系、华夏汉语的发生和发展、人类语言的共性和语言法则的实体、人脑语言的电脑移植和语言基因图谱分析、东西方语言文化的本原差异及其分野、世界语言学理论的建构和人类语言的世界化进程……，其旨趣在于对语言学"公理"逐一独立思考。一些研究结果曾经先后以论文的形式在海内外发表，得到海内外一些学人的关注，其中题名《论索绪尔语言符号任意性原则的失误与复归》(1994)的论文引发《语言文字应用》编辑部主持的90年代语言观的讨论。这一系列探索中，贯穿着一根主线：所谓人性就是语言性，人与动物的区别就在于创造并运用语言符号以顽强地表现自己的精神世界，语言是人类的精神家园。

语言是生物的、心理的、物理的、社会的、艺术的……。语言的研究或语言学，归根结底就是人性(生物人、心理人)和人学(艺术人、社会人、经济人、科技人)的研究，作为自然界的精灵，人类是人文(最初是文身)、社会(最初是血族)和科学science(最初是巫术)的生物。人类的使命就是"认识你自己"，认识内心中的感受，认识社会界的关系，认识自然界的奥秘，由此形成人文

学科、社会学科和自然学科的分野。而作为认知手段的系统和人类文化的底座，语言符号系统是贯通人文、社会和自然的基本工具，语言学是渗透于人文、社会和自然一切理性认知或艺术表现领域的泛学科。语言的研究方法或语言学的理论观点，应当是一切学科的研究方法及其理论观点的融会贯通。换而言之，人文学科、社会学科和自然学科的畛域在语言学中趋于消解和融合。

自从古希腊的"哲学·语言学·自然学"的原初学术混沌母体的解构以后，人文主义与科学主义就成为欧洲学术史上争执不休的话题，在近代工业文明的背景上，伴随着欧洲文明的世界性洪溢和物质时尚的挡不住诱惑而科学主义甚嚣尘上。中国传统学术主张"天人合一"，因此没有出现人文与科学的壁垒森严，而是包容为"道"与"器"的合二而一，在几千年的历史中具体表现为宗法伦理的"和"和日常技术的"用"。随着近现代东西文化的撞击和学术研究的渗透，西方人文主义与科学主义的纷争以东西方文化类型的歧异方式转移到现代中国，成为中国现当代文化的论题。

在史无前例的文化革命结束以后，伴随着生存环境的解冻和基本人性的复苏，20 世纪 80 年代的中国学术界曾经掀起过一阵子文化热。根据个人秉性，在"热"中，总是沉默的。1989 年，中国语言学界发生过人文主义与科学主义的针锋相对。与申小龙主张汉语研究的人文主义本体论相反，朱晓农认为"中国语言学的'科学主义'不是'宠坏了的科学主义'，而是远未到家的科学主义"[①]，在朱晓农的相关论文中再三强调中国现代语言

① 申小龙、朱晓农《关于语言本体论和语言学方法论的对话》，《北方论丛》1989 年第 1 期。

学最为紧缺的是科学主义的精神、理性主义的认识论和演绎主义的方法论。申小龙强调既成汉语研究没有植根于汉语事实，对汉语的个性特征务必突显，朱晓农提醒汉语语言学研究落后于当代西方语言学，西方学术精神和方法务必重视。平心而论，可谓各有千秋。虽然如此，然而在科学主义与人文主义之争中，有几个概念值得深入思考：

1.语言学是“科学”(science)还是“学科”(subject )？语言学是否与“自然科学”的学科性质一样？2.科学精神是不是等同于科学主义？人文精神是不是等同于人文主义？理性认识论与非理性认识论，自然科学认识论与其他学科认识论，演绎方法论与非演绎方法论，存在着怎样的关系？

首先，随着文字的发明而进入文明阶段，人类对语言进行自觉分析的本原语言文字学是一切学科之母。从人类学科体系来看，语言学既不隶属于“自然科学”，也不属于社会学科或人文学科，语言学是一门不同于也不应隶属于任何其他学科的学科，然而是所有一切学科底座的特殊学科，其关系大致如下：

| | | |
|---|---|---|
| …音乐、舞蹈、美术、文学… | 宗教、权术、历史、法律、经济… | 数学、物理、化学、生物… |
| ……人文学科…… | ……社会学科…… | ……自然学科…… |
| 艺术……哲学……物理学 | | |
| 学科之母——语言文字学 | | |
| 神话·语言·巫术 | | |
| 非语言文字符号 | | |
| 对象世界存在……人类感受认知……生物大脑机制 | | |

其次,无论什么精神或方法,一旦冠上"主义"则难免染上独尊性或排他性。对于语言学来说,科学性和人文性、理性和感性、演绎性和归纳性、描写性和阐释性,天然存在着互补和交叉关系。从客观对象的认知来说,科学性和人文性都具有主体性,理性和感性都具有认知性,演绎性和归纳性都具有逻辑性,描写性和阐释性都具有陈述性。作为意识形态和客观世界中介的符号系统,必然具有复杂性和综合性,因此语言学离不开各种方法的借鉴,但是绝对没有一种"唯我独尊"的语言学精神和方法。归根结底,既然语言学是一切学科之母和一切学科的底座,语言学方法也就是一切学科方法的基础和综合。现代科学已经证明,所谓最精密的数学也具有模糊性,所谓可实验性的物理也随着观察的主体而具有差异性。对于语言学来说,学术研究需要科学精神——建立在可证伪性基础上的怀疑或批判精神,但是绝不需要盲目或偏执的"科学主义"。

在汉语中,作为外来词的"赛先生"至少具有两个含义:1.作为学科或知识体系的"科学",名词,可以指"自然科学"(科学院),也可以指"学科"或"学术体系"(社会科学院);2.作为真理性评判标准的"科学",形容词,如"不/很科学"。

"科学精神"是追求理解的认知精神和一定的学术规范,而"科学主义"是以学科现有结论或规范作为评判的惟一选择。然而,作为学科的科学绝对不等同真理,其中虽然包含部分相对真理,即现在认可、未被证伪的结论或假说,同时也包含已被证伪而仍然被某些人认可的一些论点。科学的本质是假说,科学的生命在于可证伪性,科学的历史就是一部学科研究结论的不断证伪史。只有宗教才自认为不具有可证伪性,教义是无可质疑的绝对真理。换而言之,作为一种异己力量的"科学主义",本质上背离了从古希腊就开始的注重探索性和批判性的科学精神,

否定了"认识你自己"的理性精神。

20世纪的中国现代语言学，为什么总是跟着西洋语言学一步一趋？中国语言学滞后的原因既不在于科学主义的障碍，也不在于人文主义的欠缺，而在于非批判性借鉴或机械模仿必然导致的西方理论框架和汉语语言事实的错位。西方理论框架并非贯穿的就是科学主义，结构主义的哲学基础是社会心理学，生成语言学的哲学基础是唯理主义，情景语言学的哲学基础是人类文化学。框架和事实错位的结果提醒我们，中国语言学滞后的症结是东西方语言学合流基点的确定错误。东西方语言学合流的基点不应是欧美语具有而汉语所无的语形语法学框架，而应是汉语和欧美语皆有的语义结构本位，因为语言的本质共性是语义性。

根据埃德蒙·利奇(Edmund Leach)的观点，与其说社会科学(未单列人文科学，可能包括在其内)与自然科学这两大领域是相同的领域，不如说只是相似的领域。社会科学资料的内在特征，如人类社会在一定程度上有意志、文化事实难以分为一些能够精确分类描述的基本单位等，决定了把自然科学的方法借用到社会科学中大可怀疑[1]。虽然社会科学与自然科学只是相似的领域，以自然科学方法替代社会科学方法此路不通，但是不等于自然科学方法在社会科学中就寸步难行，二者之间毫无共通之处。虽然，我并不反对在"科学主义"或违背了科学精神的"科学主义思潮"招摇过市的今天，务必泼些冷水。因为以研究自然事物的冷漠态度和刻板方式，去研究或管理具有情感、意志、思想和表现力的人类，注定难免碰壁。如果我们把"人文"扩展为广义，从根本上来说，科学也是一种人文，是人的认知能力

[1] 埃德蒙·利奇《文化与交流》(郭凡等译)，上海人民出版社2000年。

对自然对象的认知。

然而,问题的最终症结在于——在人类认识论上往往存在一个共同的盲点:以为"科学思维"是人类认识世界的惟一手段,以为所有"学科"都建筑在相同的或惟一的认知方式上。从人类意识形态演变的连续统来说,人类的认知方式大致有五种:

1. 以感觉和印象为特征的生物本能方式——日常行为;
2. 以心象和镜象为特征的艺术表现方式——人文艺术;
3. 以像似和隐喻为特征的认知符号方式——语言符号;
4. 以谋划和规范为特征的学术体系方式——社会学科;
5. 以实验和运算为特征的学术体系方式——自然学科。

只有认识到以像似和隐喻为特征的语言符号方式是在生物本能方式和艺术表现方式的基础上成长起来的,又以此为基点孕育出社会学科和自然学科的学术体系方式,同时以语言为载体形成文学艺术以及相关人文艺术学科。只有认识到语言这一认知符号系统中的感性因素和理性因素的渗透,只有认识到语言符号中的人文性、社会性和科学性因素的交融,才能够最终走出"科学主义"与"人文主义"相互对立的怪圈。只有确立语言的本体是人类的认知行为、符号的结构是对象的结构临摹,语言的表达是内心的感受宣泄,语言学的研究就是揭示语言行为的认知性、语言结构的像似性和语言表达的宣泄性,才能够拆除普罗克拉斯提斯之床。

当今学术潮流浩浩荡荡,正向多元、渗透、互补、宽容方向澎湃奔流,此拙著名之《理论语言学:人文与科学的双重精神》,愿学界同好不吝赐教。

东亭西溪　李葆嘉谨识

2001年6月5日于随园

对自然对象的认知。

然而,问题的最终症结在于——在人类认识论上往往存在一个共同的盲点:以为"科学思维"是人类认识世界的惟一手段,以为所有"学科"都建筑在相同的或惟一的认知方式上。从人类意识形态演变的连续统来说,人类的认知方式大致有五种:

1. 以感觉和印象为特征的生物本能方式——日常行为;

2. 以心象和镜象为特征的艺术表现方式——人文艺术;

3. 以像似和隐喻为特征的认知符号方式——语言符号;

4. 以谋划和规范为特征的学术体系方式——社会学科;

5. 以实验和运算为特征的学术体系方式——自然学科。

只有认识到以像似和隐喻为特征的语言符号方式是在生物本能方式和艺术表现方式的基础上成长起来的,又以此为基点孕育出社会学科和自然学科的学术体系方式,同时以语言为载体形成文学艺术以及相关人文艺术学科。只有认识到语言这一认知符号系统中的感性因素和理性因素的渗透,只有认识到语言符号中的人文性、社会性和科学性因素的交融,才能够最终走出"科学主义"与"人文主义"相互对立的怪圈。只有确立语言的本体是人类的认知行为、符号的结构是对象的结构临摹,语言的表达是内心的感受宣泄,语言学的研究就是揭示语言行为的认知性、语言结构的像似性和语言表达的宣泄性,才能够拆除普罗克拉斯提斯之床。

当今学术潮流浩浩荡荡,正向多元、渗透、互补、宽容方向澎湃奔流,此拙著名之《理论语言学:人文与科学的双重精神》,愿学界同好不吝赐教。

东亭西溪　李葆嘉谨识

2001年6月5日于随园